

Founded in 1929 by John Dewey and Carl Murchison

THE JOURNAL OF SOCIAL PSYCHOLOGY

ERNEST BEAGLEHOLE
Victoria University Col-
lege, New Zealand

S. BIESHEUVEL
National Institute for
Personnel Research,
Johannesburg

HAROLD E. BURTT
Ohio State University

DORWIN CARTWRIGHT
University of Michigan

RAYMOND B. CATTELL
University of Illinois

KAMLA CHOWDHRY
Ahmedabad Textile Indus.
Res. Assoc., India

RICHARD S. CRUTCHFIELD
University of California

A. H. EL-KOUSSY
Ein Shams University

OTTO KLINEBERG
University of Paris

I. D. MACCRONE
University of the
Witwaterstrand

R. B. MACLEOD
Cornell University

EMILIO MIRA Y LOPEZ
Rio de Janeiro

If this space should be unstamped, this is the regular library edition.
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not
be resold for a period of five years from date of publication.

GARDNER MURPHY
Menninger Foundation

O. A. OESER
University of Mel-
bourne, Australia

A. T. POFFENBERGER
Columbia University

E. TERRY PROTHRO
American University of
Beirut, Lebanon

BERTRAND RUSSELL
Petersfield, England

KOJI SATO
Kyoto University

MARSHALL H. SEGALL
State University
of Iowa

SADAJI TAKAGI
Tokyo University

WARREN H. TEICHER
Tufts University

GOODWIN WATSON
Columbia University

KIMBALL YOUNG
Northwestern University

PAUL THOMAS YOUNG
University of Illinois

VOLUME 67

1965

Copyright © 1965 by The Journal Press
Provincetown, Massachusetts, U. S. A.

All rights reserved

Published at 35 New St., Worcester, Mass.
Second-class postage paid at Worcester, Mass.

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts

AUTHOR INDEX

Bales, Robert F.	239	Lindgren, Henry Clay	23
Binner, Marian Roco	115	Long, Barbara H.	139
Boruchow, Joan K.	61	Maglione, Frank D.	329
Brandon, Arlene C.	343	McCann, Brian M.	227
Carr, Lester	259	McConville, Carolyn B.	45
Cattell, Raymond B.	77	McDavid, John	115
Comrey, Andrew L.	357	Mishra, H. K.	197
Cooley, C. Ewing	325	Murgatroyd, Dorothy	201
Das, J. P.	197	Nanda, P. C.	197
Denmark, Florence	201	Newmeyer, John A.	357
Diab, Lutfy N. 283, 297,	303	Papageorgis, Demetrios	227
Diers, Carol Jean	97	Pepitone, Albert	201
Doob, Leonard W.	3	Plant, Walter T.	67
Drinkwater, Barbara L.	189	Proenza, Luis	273
Gaier, Eugene L.	379	Rabow, Jerome	31
Garbutt, John T.	269	Roberts, S. Oliver	259
Genskow, Jack K.	329	Rosnow, Ralph L. 39,	211
Gursuch, Richard L.	77	Sampson, Edward E.	343
Hare, A. Paul	239	Smart, Reginald G.	221
Hartley, Raymond E.	149	Stimpson, David Vern	105
Hemphill, John K.	45	Stotland, Ezra	315
Hudson, Bradford B.	115	Strickland, Bonnie R.	273
Hutton, Jerry B.	325	Telford, Charles W.	67
Knapp, Robert H. 269,	371	Thomas, Joseph A.	67
Krause, Merton S.	163	White, William F.	379
Lindgren, Fredrica	23	Ziller, Robert C.	139

TABLE OF CONTENTS

Exploring eidetic imagery among the Kamba of central Kenya	3
By LEONARD W. DOOB	
Creativity, brainstorming, and orneriness: A cross-cultural study . . .	23
By HENRY CLAY LINDGREN AND FREDRICA LINDGREN	
Quantitative aspects of the group-psychotherapist's role behavior: A method- ological note	31
By JEROME RABOW	
A delay-of-reinforcement effect in persuasive communication	39
By RALPH L. ROSNOW	
The effect of "human" vs. "machine" set on group problem-solving procedures .	45
By JOHN K. HEMPHILL AND CAROLYN B. McCONVILLE	
Characteristics of nurses in short- and long-term treatment settings . . .	61
By JOAN K. BORUCHOW	
Some personality differences between dogmatic and nondogmatic groups . .	67
By WALTER T. PLANT, CHARLES W. TELFORD, AND JOSEPH A. THOMAS	
The definition and measurement of national morale and morality . . .	77
By RAYMOND B. CATTELL AND RICHARD L. GORSUCH	
Social-desirability ratings of personality items by three subcultural groups .	97
By CAROL JEAN DIERS	
The effects of discrepant information on perceptual processes: An application of balance theory	105
By DAVID VERN STIMPSON	
Group responses to threat: Perception and learning under stress . . .	115
By BRADFORD B. HUDSON, JOHN McDAVID, AND MARIAN ROCO BINNER	
Some correlates of the don't-know response in opinion questionnaires . .	139
By ROBERT C. ZILLER AND BARBARA H. LONG	
Perceptual accentuation as a problem in the psychology of judgment . . .	149
By RAYMOND E. HARTLEY	
Role-deviant respondent sets and resulting bias, their detection and control in the survey interview	163
By MERTON S. KRAUSE	
BOOKS	185
BOOKS RECENTLY RECEIVED	185
A comparison of the direction-of-perception technique with the Likert method in the measurement of attitudes	189
By BARBARA L. DRINKWATER	
Discrimination of geometrical patterns in tribal, rural, and urban children .	197
By P. C. NANDA, J. P. DAS, AND H. K. MISHRA	
Effect of differential valuation on group level of aspiration, decision time, and productivity	201
By FLORENCE DENMARK, DOROTHY MURGATROYD, AND ALBERT PEPITONE	

Bias in evaluating the presidential debates: A "Splinter" effect	211
By RALPH L. ROSNOW	
Social-group membership, leadership, and birth order	221
By REGINALD G. SMART	
Effect of discrepant communications on self-perception	227
By DEMETRIOS PAPAGEORGIS AND BRIAN M. MCCANN	
Diagnostic use of the interaction profile	239
By ROBERT F. BALES AND A. PAUL HARE	
Correlates of civil-rights participation	259
By LESTER CARR AND S. OLIVER ROBERTS	
Variation in time descriptions and need achievement	269
By ROBERT H. KNAPP AND JOHN T. GARBUTT	
A study of prejudice in Negro and white college students	273
By LUIS PROENZA AND BONNIE R. STRICKLAND	
Studies in social attitudes: I. Variations in latitudes of acceptance and rejection as a function of varying positions on a controversial social issue	283
By LUTFY N. DIAB	
Studies in social attitudes: II. Selectivity in mass communication media as a function of attitude-medium discrepancy	297
By LUTFY N. DIAB	
Studies in social attitudes: III. Attitude assessment through the semantic-dif- ferential technique	303
By LUTFY N. DIAB	
Experimental social psychology and its neighbors	315
By EZRA STOTLAND	
Adolescent response to religious appeal as related to IPAT anxiety	325
By C. EWING COOLEY AND JERRY B. HUTTON	
Familiarity, dogmatism, and reported student attitudes toward the disabled	329
By JACK K. GENSKOW AND FRANK D. MAGLIONE	
Demand characteristics vs. physiological needs: A methodological critique of the "new look" school of perception	343
By ARLENE C. BRANDON AND EDWARD E. SAMPSON	
Measurement of radicalism-conservatism	357
By ANDREW L. COMREY AND JOHN A. NEWMAYER	
The language of postural interpretation	371
By ROBERT H. KNAPP	
Modes of conformity and career selection of rural and urban high-school seniors	379
By EUGENE L. GAIER AND WILLIAM F. WHITE	
BOOKS RECENTLY RECEIVED	393

Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, published in March, June, September, and December. Devoted to child behavior, animal behavior, and the problems of aging. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1891 at \$15.00 per volume, plus transportation.

Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, published in January, April, July, and October. Devoted primarily to experimental, theoretical, clinical, and historical psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1927 at \$15.00 per volume, plus transportation.

Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. National, group, cultural, racial, and differential psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Bimonthly, three volumes per year; per annum, \$30.00; per volume, \$15.00; single numbers, \$7.50. First number is dated February, 1930. Complete sets from 1930 at \$15.00 per volume, plus transportation.

Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

Club Rate for Five International Psychological Journals

THE JOURNAL OF GENETIC PSYCHOLOGY	PER ANNUM
(Quarterly, two volumes per year)	\$20.00
GENETIC PSYCHOLOGY MONOGRAPHS	
(Quarterly, two volumes per year)	\$20.00
THE JOURNAL OF GENERAL PSYCHOLOGY	
(Quarterly, two volumes per year)	\$20.00
THE JOURNAL OF SOCIAL PSYCHOLOGY	
(Bimonthly, three volumes per year)	\$30.00
THE JOURNAL OF PSYCHOLOGY	
(Bimonthly, three volumes per year)	\$30.00
TOTAL	\$120.00
TEACHERS AND STUDENTS EDITION, <i>Restricted Sale</i>	\$35.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly: February, May, August, November. Two volumes per year. Child behavior, animal behavior, and the problems of aging. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1926 at \$15.00 per volume, plus transportation.

NOTICE. Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology*. Our policy remains unchanged for regular monographs: Authors will be charged for cuts and expensive tables, but there will be no page charge for the first 50 pages. For all pages above 50 there will be a charge of \$10.00 per page. Two hundred reprints of each monograph will be supplied free of charge, but there will be nominal charges for such things as vitas, special covers, or special title pages.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.

Genetic Psychology Monographs (continued)

VOLUME 1—1926

1. Performance tests for children of pre-school age—R. STUTSMAN
2. An experimental study of the eidetic type—H. KLÜVER
- 3 & 4. A study of natio-racial mental differences—N. D. M. HIRSCH
5. A psychological study of juvenile delinquency by group methods—J. W. BRIDGES AND K. M. B. BRIDGES
6. The influence of puberty praecox upon mental growth—A. GESELL

VOLUME 2—1927

- 1 & 2. The mind of a gorilla—R. M. YERKES
3. The role of eye-muscles and mouth-muscles in the expression of the emotions—K. DUNLAP
4. Family similarities in mental-test abilities—R. R. WILLOUGHBY
5. Coordination in the locomotion of infants—L. H. BURNSIDE
6. The mind of a gorilla: Part II. Mental development—R. M. YERKES

VOLUME 3—January-June, 1928

1. An experimental study of the olfactory sensitivity of the white rat—J. R. LIGGETT
2. A photographic study of eye movements in reading formulæ—M. A. TINKER
3. An experimental study of the East Kentucky mountaineers—N. D. M. HIRSCH
4. Responses of foetal guinea pigs prematurely delivered—G. T. AVERY
5. Objective differentiation between three groups in education (teachers, research workers, and administrators)—M. B. JENSEN
6. The effect of segregation on the sex behavior of the white rat as measured by the obstruction method—M. JENKINS

VOLUME 4—July-December, 1928

1. Observation and training of fundamental habits in young children—E. A. BOTT, W. E. BLATZ, N. CHANT, AND H. BOTT
- 2 & 3. Determination of a content of the course in literature of a suitable difficulty for junior and senior high school students—M. C. BURCH
- 4 & 5. Methods for diagnosis and treatment of cases of reading disability—M. MONROE
6. The relative effectiveness of lecture and individual reading as methods of college teaching—E. B. GREENE

VOLUME 5—January-June, 1929

1. The age factor in animal learning: I. Rats in the problem box and the maze—C. P. STONE
2. The effect of delayed incentive on the hunger drive in the white rat—E. L. HAMILTON
3. Which hand is the eye of the blind?—J. M. SMITH
4. The effect of attitude on free word association-time—A. G. EKDAHL
5. The localization of tactual space: A study of average and constant errors under different types of localization—L. E. COLE
6. The effects of gonadectomy, vasotomy, and injections of placental and orchic extracts on the sex behavior of the white rat—H. W. NISSEN

VOLUME 6—July-December, 1929

1. Learning and growth in identical infant twins: An experimental study by the method of co-twin control—A. GESELL AND H. THOMPSON
2. The age factor in animal learning: II. Rats on a multiple light discrimination box and a difficult maze—C. P. STONE
3. The acquisition and interference of motor habits in young children—E. MCGINNIS
4. A vocational and socio-educational survey of graduates and non-graduates of small high schools of New England—A. D. MUELLER
- 5 & 6. A study of the smiling and laughing of infants in the first year of life—R. W. WASHBURN

VOLUME 7—January-June, 1930

1. Tensions and emotional factors in reaction—E. DUFFY
2. Teacher influence on class achievement: A study of the relationship of estimated teaching ability to pupil achievement in reading and arithmetic—H. R. TAYLOR
- 3 & 4. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior—P. H. EWERT
5. A study of the mental development of children with lesion in the central nervous system—E. E. LORD
6. An experimental study upon three hundred school children over a six-year period—N. D. M. HIRSCH

VOLUME 8—July-December, 1930

1. The amount and nature of activities of newborn infants under constant external stimulating conditions during the first ten days of life—O. C. IRWIN
2. Race and social differences in performance tests—S. D. PORTEUS, *et al.*
3. Language and growth: The relative efficacy of early and deferred vocabulary training, studied by the method of co-twin control—L. C. STRAYER
4. Eye-movements and optic nystagmus in early infancy—J. M. MCGINNIS
- 5 & 6. Reactions of kindergarten, first-, and second-grade children to constructive play materials—L. FARWELL

VOLUME 9—January-June, 1931

- 1 & 2. The status of the first-born with special reference to intelligence—H. H. HSIAO
- 3 & 4. An experimental study of bright, average, and dull children at the four-year mental level—H. P. DAVIDSON
5. An historical, critical, and experimental study of the Seashore-Kwalwasser test battery—P. R. FARNSWORTH
6. A comparison of difficulty and improvement in the learning of bright and dull children in reproducing a descriptive selection—F. T. WILSON

VOLUME 10—July-December, 1931

1. A comparative study of a group of southern white and negro infants—M. B. MCGRAW
- 2 & 3. An experimental study of prehension in infants by means of systematic cinema records—H. M. HALVERSON
4. The limits of learning ability in kittens—A. M. SHUEY
- 5 & 6. The effect of habit interference upon performance in maze learning—O. W. ALM

VOLUME 11—January-June, 1932

1. General factors in transfer of training in the white rat—T. A. JACKSON
2. The effect of color on visual apprehension and perception—M. A. TINKER
3. The reliability and validity of maze experiments with white rats—R. LEEPER
4. A critical study of two lists of best books for children—F. K. SHUTTLEWORTH
- 5 & 6. Measuring human energy cost in industry: A general guide to the literature—R. M. PAGE

VOLUME 12—July-December, 1932

1. Family resemblances in verbal and numerical abilities—H. D. CARTER
2. The development of fine prehension in infancy—B. M. CASTNER
- 3 & 4. The growth of adaptive behavior in infants: An experimental study at seven age levels—H. M. RICHARDSON
- 5 & 6. Differential reactions to taste and temperature stimuli in newborn infants—K. JENSEN

VOLUME 13—January-June, 1933

1. A critique of sublimation in males: A study of forty superior single men—W. S. TAYLOR
2. A study of the nature, measurement, and determination of hand preference—H. L. KOCH, *et al.*
3. The growth and decline of intelligence: A study of a homogeneous group between the ages of ten and sixty—H. E. JONES AND H. S. CONRAD
4. The relation between the complexity of the habit to be acquired and the form of the learning curve in young children—M. L. MATTSOHN
5. Eating habits in relation to personality development of two- and three-year-old children: A study of sixty-nine children in two nursery schools—A. A. ELIOT
6. Coordinating mechanisms of the spinal cord—O. C. INGEBRITSSEN

Genetic Psychology Monographs (continued)

VOLUME 14—July-December, 1933

1. Mental growth during the first three years: A developmental study of sixty-one children by repeated tests—N. BAYLEY
2. A study of triplets: including theories of their possible genetic relationships—F. N. ANDERSON AND N. V. SCHEIDEMANN
3. The objective measurement of emotional reactions—H. V. GASKILL
4. Development of behavior in the fetal cat—J. D. CORONIOS
5. A study of certain language developments of children in grades four to twelve, inclusive—L. L. LABRANT
6. The effect of early and delayed practice on memory and motor performances studied by the method of co-twin control—J. R. HILGARD

VOLUME 15—January-June, 1934

1. Studies in the psychology of tone and music—P. R. FARNSWORTH
2. Motor learning of children in equilibrium in relation to nutrition—E. L. BEEBE
3. Discrimination limits of pattern and size in the goldfish *Carassius auratus*—J. B. ROWLEY
4. Limits of learning ability in the white rat and the guinea pig—B. F. RIESS
- 5 & 6. The limits of learning ability in rhesus monkeys—H. A. FJELD

VOLUME 16—July-December, 1934

1. A statistical study of ratings on the California Behavior Inventory for Nursery-School Children—H. S. CONRAD
2. An eye-movement study of objective examination questions—A. FRANDSEN
3. An experimental study of constitutional types—O. KLINEBERG, S. E. ASCH, AND H. BLOCK
4. The development of a battery of objective group tests of manual laterality, with the results of their application to 1300 children—W. N. DUNST
- 5 & 6. An experimental study in the prenatal guinea-pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life—L. CARMICHAEL

VOLUME 17—January-December, 1935

1. Organization of behavior in the albino rat—R. L. THORNDIKE
2. Brightness discrimination in the rhesus monkey—M. P. CRAWFORD
3. The limits of learning ability in cebus monkeys—A. M. KOCH
4. Nature-nurture and intelligence—A. M. LEAHY
5. On intelligence of epileptic children—E. B. SULLIVAN AND L. GAHAGAN
6. A study of the play of children of preschool age by an unobserved observer—D. L. COCKRELL

VOLUME 18—January-December, 1936

1. Sex differences in variational tendency—Q. McNEMAR AND L. M. TERMAN
2. The process of learning to dress among nursery-school children—C. B. KEY, M. R. WHITE, M. P. HONZIK, A. B. HEINEY, AND D. ERWIN
3. A study of the present social status of a group of adults, who, when they were in elementary schools, were classified as mentally deficient—W. R. BALLER
4. The influence of specific experience upon mental organizations—A. ANASTASI
- 5 & 6. Studies in aggressiveness—L. BENDER, S. KEISER, AND P. SCHILDER

VOLUME 19—January-December, 1937

1. Psychological bases of self-mutilation—C. DABROWSKI
Masculine temperament and secondary sex characteristics: A study of the relationship between psychological and physical measures of masculinity—H. GILKINSON
2. A psychological study of forty unmarried mothers—R. D. NOTTINGHAM
3. Behavior problems in the children of psychotic and criminal parents—L. BENDER
4. The sequential patterning of prone progression in the human infant—L. B. AMES

VOLUME 20—January-December, 1938

1. The relationship between characteristics of personality and physique in adolescents—P. S. DE Q. CAROT
2. Behavior problems of elementary school children: A descriptive and comparative study—I. Y. MASTEN
3. Graphic representation of a man by four-year-old children in nine prescribed drawing situations—P. F. GRIDLEY
4. Differences between two groups of adult criminals—R. S. TOLMAN
4. A comparative study by means of the Rorschach method of personality development in twenty pairs of identical twins—E. THOMP

Individual differences in the facial expressive behavior of preschool children: A study by the time-sampling method—C. SWAN

VOLUME 21—January-December, 1939

1. An experimental analysis of "level of aspiration"—R. GOULD
2. Some light on the problem of bilingualism as found from a study of the progress in mastery of English among pre-school children of non-American ancestry in Hawaii—M. E. SMITH
3. Domination and social integration in the behavior of kindergarten children and teachers—H. H. ANDERSON
3. The capacity of the rhesus and cebus monkey and the gibbon to acquire differential response to complex visual stimuli—W. E. GALT
4. The social-sex development of children—E. H. CAMPBELL

VOLUME 22—January-December, 1940

1. Measuring human relations: An introduction to the study of the interaction of individuals—E. D. CHAPPEL
2. Aggressive behavior in young children and children's attitudes toward aggression—M. D. FITE
3. Student attitudes toward religion—E. NELSON
3. The prediction of the outcome-on-furlough of dementia praecox patients—J. S. JACOB
4. Significant characteristics of preschool children as located in the Conrad inventory—K. H. READ
4. Learning by children at noon-meal in a nursery school: Ten "good" eaters and ten "poor" eaters—J. B. MCCAY, E. B. WARING, AND P. J. KRUSE

Studies in the interpretation of play: I. Clinical observation of play disruption in young children—E. H. ERIKSON

VOLUME 23—January-June, 1941

1. An analysis of certain variables in a developmental study of language—F. M. YOUNG
1. Infant development under conditions of restricted practice and of minimum social stimulation—W. DENNIS
2. An analysis of the mental factors of various age groups from nine to sixty—B. BALINSKY
2. Factors influencing performance on group and individual tests of intelligence: I. Rate of work—M. W. BENNETT
- Individual differences in apperceptive reaction: A study of the response of preschool children to pictures—E. W. AMEN

VOLUME 24—July-December, 1941

1. Twins T and C from infancy to adolescence: A biogenetic study of individual differences by the method of co-twin control—A. GESSELL AND H. THOMPSON
1. Finger nail-biting: Its incipency, incidence, and amelioration—A. L. BILLIC
1. An experimental study of the factors of maturation and practice in the behavioral development of the embryo of the frog, *Rana pipiens*—A. FROMME
2. The Fels child behavior scales—T. W. RICHARDS AND M. P. SIMONS
2. Measurement of the size of general English vocabulary through the elementary grades and high school—M. K. SMITH
2. Stereotypes in the field of musical eminence—P. R. FARNSWORTH

VOLUME 25—January-June, 1942

1. A study of factors determining family size in a selected professional group—J. C. FLANAGAN
1. A genetic study of geometrical-optical illusions—A. WALTERS
1. Interpretation of behavior-ratings in terms of favorable and unfavorable deviations: A study of scores from the Read-Conrad Behavior Inventory—K. H. READ AND H. S. CONRAD
2. Are there any innate behavior tendencies?—J. B. SCHOELLAND
2. An investigation of the intelligibility of the speech of the deaf—C. V. HUDGINS AND F. C. NUMBERS

Genetic Psychology Monographs (continued)

VOLUME 26—July-December, 1942

1. The critical frequency limen for visual flicker in children between the ages of 6 and 18—V. L. MILLER
Some factors determining handedness in the white rat—K. L. WENTWORTH
2. Motivation and behavior—E. FRENKEL-BRUNSWIK

VOLUME 27—January-June, 1943

1. Comparison of children's personality traits, attitudes, and intelligence with parental occupation—N. R. MADDY
2. A comparative study of mental functioning patterns of problem and non-problem children seven, eight, and nine years of age—M. L. PIGNATELLI

VOLUME 28—July-December, 1943

1. Separation anxiety in young children: A study of hospital cases—H. EDLSTON
2. Correlates of vocational preferences—W. A. BRADLEY, JR.

VOLUME 29—January-June, 1944

1. Mental changes after bilateral prefrontal lobotomy—S. D. PORTEUS AND R. D. KEPNER
2. A twin-controlled experiment on the learning of auxiliary languages—B. PRICE, W. J. KOSTER, AND W. M. TAYLOR

VOLUME 30—July-December, 1944

1. A method of administering and evaluating the thematic appreciation test in group situations—R. M. CLARK
2. A study of anxiety reactions in young children by means of a projective technique—R. TEMPLE AND E. W. AMEN

VOLUME 31—January-June, 1945

1. The evolution of intelligent behavior in rhesus monkeys—B. WEINSTEIN
2. Perceptual behavior of brain-injured, mentally defective children: An experimental study by means of the Rorschach technique—H. WERNER

VOLUME 32—July-December, 1945

1. A clinical study of sentiments: I.—H. A. MURRAY AND C. D. MORGAN
2. A clinical study of sentiments: II.—H. A. MURRAY AND C. D. MORGAN

VOLUME 33—January-June, 1946

1. Interpretation of spontaneous drawings and paintings—T. S. WAERNER
Preferences for sex symbols and their personality correlates—K. FRANCK
2. Outstanding traits: In a selected college group, with some reference to career interests and war records—F. L. WELLS AND W. L. WOODS

VOLUME 34—July-December, 1946

1. The relation of emotional adjustment to intellectual function—J. L. DESPERT AND H. O. PIERCE
2. The smiling response: A contribution to the ontogenesis of social relations—R. A. SPITZ
3. Finger-painting and personality diagnosis—P. J. NAPOLI

VOLUME 35—January-June, 1947

1. The thematic apperception technique in the study of culture-personality relations—W. E. HENRY
2. A continuation study of anxiety reactions in young children by means of a projective technique—M. DORKEY AND E. W. AMEN

VOLUME 36—July-December, 1947

1. Maze test validation and psychosurgery—S. D. PORTEUS AND H. N. PETERS
2. The diagnostic implications of Rorschach's test in case studies of mental defectives—I. JOLLES

VOLUME 37—January-June, 1948

1. The radio day time serial: A symbolic analysis—W. L. WARNER AND W. E. HENRY
2. The relation of personality characteristics and response to verbal approval in a learning task—G. L. GRACE
3. The mechanism of vision: XVIII. Effects of destroying the visual "associative areas" of the monkey—K. S. LASHLEY
4. A study of the relationship between handwriting and personality variables—P. CASTELNUOVA-TEDESCO

VOLUME 38—July-December, 1948

1. Modern language learning: The intensive course as sponsored by the United States Army, and implications for the undergraduate course of study—M. LIND
2. Conflict: A study of some interactions between appetite and aversion in the white rat—M. A. TOLCOTT
3. Schizophrenia and the MAPS test: A study of certain formal psycho-social aspects of fantasy production in schizophrenia as revealed by performance on the Make a Picture Story (MAPS) Test—E. S. SHNEIDMAN
4. A study of the transmission of authority patterns in the family—H. L. INGERSOLL

VOLUME 39—January-June, 1949

1. A study of the psychoanalytic theory of psychosexual development—G. S. BLUM
2. The assessment of parental attitudes in relation to child adjustment—E. J. SHOEN, JR.
3. Qualitative differences in the vocabulary responses of normals and abnormals—H. FEIFEL
4. The relative effectiveness of motion and still pictures as stimuli for eliciting fantasy stories about adolescent-parent relationships—P. E. EISERER
5. The organization of hereditary maze-brightness and maze-dullness—L. V. SEARLE

VOLUME 40—July-December, 1949

1. An experimental study of what young school children expect from their teachers—B. BIBER AND C. LEWIS
2. A study of the relative effects of age and of test difficulty upon factor patterns—H. A. CURTIS
3. A projective experiment using incomplete stories with multiple choice endings—J. K. SEATON
4. Effects of sex role and social status on the early adolescent personality—E. MILNER
5. Social perceptions and attitudes of children—M. RADKE, H. TRACER, AND H. DAVIS

VOLUME 41—January-June, 1950

1. Some psychological and educational aspects of pediatric practice: A study of well-baby clinics—L. H. BLUM
2. One-trial learning in the domestic rat—B. B. HUDSON
3. An introduction to the principles of scientific psychoanalysis—A. ELLIS
4. Awareness of racial differences by preschool children in Hawaii—D. V. SPRINGER
5. Age trends in children's evaluation of teacher-approved and teacher-disapproved behavior—S. L. WITRYOL
6. The relationship between level of vocational aspiration and certain personal data: A study of some traits and influences bearing on the prestige level of vocational choice—J. STUBBINS

VOLUME 42—July-December, 1950

1. Personality patterns of suicidal mental hospital patients—N. L. FARBROW
2. Sex-role identification in young children in two diverse social groups—M. RABBAN
3. A study of the influence of the social field on individual behavior: As revealed in the expression of hostility and warmth by neurotics and paranoid schizophrenics in discussion group situations—D. SHAPIRO
4. An experimental study of avoidance—R. F. HEFFERLINE

VOLUME 43—January-June, 1951

1. A study of copying ability in children—E. A. TOWNSEND
2. Prestige motivation of gifted children—D. P. AUSUBEL
3. A psychological study of physical scientists—A. ROE

VOLUME 44—July-December, 1951

1. The organization of hostility controls in various personality structures—S. FISHER AND E. HINDS
2. Children and radio: A study of listeners and non-listeners to various types of radio programs in terms of selected ability, attitude, and behavior measures—E. A. RICCIUTI
3. Quantitative expression in young children—W. E. MARTIN
4. The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant—A. W. LYNNE

VOLUME 45—January-June, 1952

1. Japanese-American personality and acculturation—W. CAUDILL
2. A statistical study of the Freudian theory of levels of psychosexual development—C. A. BARNES
3. Personality characteristics of selected disability groups—D. N. WIENER

Genetic Psychology Monographs (continued)

VOLUME 46—July-December, 1952

1. The relationship of social status, intelligence, and sex of ten- and eleven-year-old children to an awareness of poverty—F. J. ESTYAN
An empirical study of the castration and Oedipus complexes—S. M. FRIEDMAN
2. The relationship between projective test scoring categories and activity preferences—M. M. SCHWARTZ
A comparison of formal and content factors in the diagnostic testing of schizophrenia—M. SHERMAN

VOLUME 47—January-June, 1953

1. Ability and accomplishment of persons earlier judged mentally deficient—D. C. CHARLES
Variations in the consistency of the behavioral meaning of personality test scores—M. KORNREICH
2. Some child-rearing antecedents of aggression and dependency in young children—R. R. SEARS, *et al*
Symptom correlates for descriptive diagnosis—J. R. WITTENBORN, *et al*.

VOLUME 48—July-December, 1953

1. Age and mental abilities: A longitudinal study—W. A. OWENS, JR.
The development of a personality questionnaire for drinkers—P. J. HAMPTON
Personality and physical disease: A test of the Dunbar hypothesis applied to diabetes mellitus and rheumatic fever—D. H. CROWELL
2. Socio-economic contrasts in children's peer culture prestige values—B. POPE
A critical review of the stability of social acceptability scores obtained with the partial-rank-order and the paired-comparison scales—S. A. WITRYOL AND G. G. THOMPSON
A study of the effects of color on Rorschach responses—G. C. BRODY

VOLUME 49—January-June, 1954

1. Factors underlying major reading disabilities at the college level—J. A. HOLMES
Parent behavior toward first and second children—J. K. LASKO
2. Social-status and intelligence: An experimental study of certain cultural determinants of measured intelligence—E. A. HAGGARD
Certain determinants and correlates of authoritarianism—S. SIEGEL
Personalities in faces: I. An experiment in social perceiving—P. F. SECORD, W. F. DUKES, AND W. BEVAN

VOLUME 50—July-December, 1954

1. A study of the relationship between play patterns and anxiety in young children—E. W. AMEN AND N. RENISON
Operational exploration of the conceptual self system and of the interaction between frames of reference—M. EDELSON AND A. E. JONES
Problem solving: A statistical description of some relationships between organismic factors and selected response measures—N. A. FATTU, E. KAPOV, AND E. V. MECH
2. The relation of cortical potentials to perceptual functions—C. CHYATTE
The import for clinical psychology of the use of tests derived from theories about infantile sexuality and adult character—D. W. MILES
Measuring personality in developmental terms: The Personal Preference Scale—M. H. KROUT AND J. K. TABIN

VOLUME 51—January-June, 1955

1. Some relations between techniques of feeding and training during infancy and certain behavior in childhood—A. BERNSTEIN
The expression of personality in drawings and paintings—L. H. STEWART
2. Negative stereotypes concerning Americans among American-born children receiving various types of minority-group education—J. A. FISHMAN
The Lincoln-Oseretsky Motor Development Scale—W. SLOAN

VOLUME 52—July-December, 1955

1. Some personality correlates of sex, sibling position, and sex of sibling among five- and six-year-old children—H. L. KOCH
A quantitative Rorschach assessment of maladjustment and rigidity in acculturating Japanese Americans—G. DEVOS
Measurement of authoritarianism and its relation to teachers' classroom behavior—H. M. MCGEE
2. The formal aspects of schizophrenic verbal communication—B. MIRIN
A study in an aspect of concept formation, with subnormal, average, and superior adolescents—H. N. HOFFMAN
Traumatic avoidance learning: Acquisition and extinction in dogs deprived of normal peripheral autonomic function—L. C. WYNNE AND R. L. SOLOMON

VOLUME 53—January-June, 1956

1. As the psychiatric aide sees his work and problems—F. L. WELLS, M. GREENBLATT, AND R. W. HYDE
An investigation of avoidance, anxiety, and escape behavior in human subjects as measured by action potentials in muscle—J. D. BROTHERS
Spread of effect: A critical review—M. H. MARX
2. Stress, fantasy, and schizophrenia: A study of the adaptive processes—O. J. B. KERNER
The attitude structure of the individual: A Q-study of the educational attitudes of professors and laymen—F. N. KERLINGER

VOLUME 54—July-December 1956

1. A study of personality differences between middle and lower class adolescents: The Szondi Test in culture-personality research—L. RAINWATER
The assessment of parental identification—S. W. GRAY AND R. KLAUS
2. The influence of social context on impulse and control tendencies in preadolescents—G. H. ZUK
Tender-mindedness versus tough-mindedness in psychology: A reexamination—H. WINTHROP
A method for the comparison of groups: A study in thematic apperception—L. C. SCHAW AND W. E. HENRY

VOLUME 55—January-June, 1957

1. Academic performance and personality adjustments of highly intelligent college students—B. M. HORRALL
The use of the Vineland Social Maturity Scale in the planning of an educational program for non-institutionalized low-grade mentally deficient children—M. G. GOTTSCHEN
2. The structure and origin of the anal character—H. BELOFF
Free expression of adolescents' interests—M. AMATORA
The rôle of mass media and the effect of aggressive film content upon children's aggressive responses and identification choices—R. S. ALBERT
Interest in persons as an aspect of sex difference in the early years—E. W. GOODENOUGH

VOLUME 56—July-December, 1957

1. Some stable response determinants of perception, thinking, and learning: A study based on the analysis of a single test—M. L. SIMMEL AND S. COUNTS
Potentials of age: An exploratory field study—S. L. PRESSEY
2. Tactile communication—L. K. FRANK
The use of a filmed puppet show as a group projective technique for children—M. R. HAWORTH
The social competence of middle-aged people—R. J. HAVIGHURST

VOLUME 57—January-June, 1958

1. Psychological and cultural problems in mental subnormality: A review of research—S. B. SARASON AND T. GLADWIN
2. Developmental aspects of discrimination in relation to adjustment—P. LONDON
Muscular tension as a response to psychological stress in rheumatoid arthritis and peptic ulcer—J. A. SOUTHWORTH
Frustration: Theory and experiment—R. LAWSON AND M. H. MARX

VOLUME 58—July-December, 1958

1. Emotional aspects of political behavior: The woman voter—E. M. BENNETT AND H. M. GOODWIN
The accuracy of self estimate: A measure of self-concept reality—R. M. BRANDT
Personality factors in social mobility: A study of occupationally mobile businessmen—J. C. ABEGLLEN
2. Interaction between authoritarian and nonauthoritarian principals and teachers—P. LAMBERT
Some factors related to the choice-status of ninety eighth-grade children in a school society—D. ELKINS
Some relationships between fantasy and self-report behavior—R. C. CALOGERAS

Genetic Psychology Monographs (continued)

VOLUME 59—January-June, 1959

1. Freshman Rorschachs and college performance—G. R. SCHMEIDLER, M. J. NELSON, AND M. BRISTOL
Perceptual aspects of attributed movement—R. W. COAN
Men and women: Personality patterns and contrasts—E. M. BENNETT AND L. R. COHEN
2. Personality factors in mothers of cerebral palsied children—G. BOLES
Separatism and integrationism: A social-psychological analysis of editorial content in New York newspapers of three American minority groups—J. A. FISHMAN AND G. S. FISHMAN
Self, role, and satisfaction—A. L. BROPHY

VOLUME 60—July-December, 1959

1. The constancy of personality ratings over two decades—R. D. TUDDENHAM
Mother-son identification and vocational interest—L. H. STEWART
A study of mother-child relationships in the emotional disorders of children—M. J. ROSENTHAL, M. FINKELSTEIN, E. NI, AND R. E. ROBERTSON
On the trail of the wolf-children—W. F. OGBURN AND N. K. BOSK
2. Measuring the mental health of normal adults—R. F. PECK
Parental identification in young children—W. EMMERICH
The problems of Jewish religious youth—B. M. LEVINSON

VOLUME 61—January-June, 1960

1. Identifying socially maladjusted school children—A. A. ADAMS
Social interaction in an interracial nursery school—H. W. STEVENSON AND N. G. STEVENSON, *et al.*
Interests of pre-adolescent boys and girls—M. AMATORA
Educational and sex differences in the organization of abilities in technical and academic students in Colombia, South America—J. F. FILKELLA
2. Some themes in the personalities of German men—L. RAINWATER
Right-left gradients in body image, body reactivity, and perception—S. FISHER
Longitudinal survey of child Rorschach responses: Younger subjects two to 10 years—L. B. AMES
Personality patterns related to occupational roles—M. SIEGELMAN AND R. F. PECK

VOLUME 62—July-December, 1960

1. The reading process: A descriptive, interdisciplinary approach—M. P. LANGMAN
Reactions to stress as a function of the level of intelligence—S. M. H. ZAIDI
Rorschach diagnosis by a systematic combining of content, thought process, and determinant scales—P. A. BOWER, R. TESTIN, AND A. ROBERTS
Longitudinal survey of child Rorschach responses: Older subjects aged 10 to 16 years—L. B. AMES
2. Twenty years of shock therapy in America, 1937-1956: An annotated bibliography—N. H. PRONKO, R. SITTERLY, AND K. BERG
The quantitative analysis of parent behavior toward psychotic children and their siblings—E. M. DONNELLY

VOLUME 63—January-June, 1961

1. Make a sentence test: An approach to objective scoring of sentence completions—E. F. BORCATT
Some aspects of the relationship between perception and motility in children—E. P. ROTHMAN
Some psychophysical aspects of mental illness in children—P. H. BERKOWITZ
Conceptual thinking in schizophrenic children—G. FRIEDMAN
2. Some maternal influences on children's personality and character—J. C. FINNEY
A study of concept formation in normals, mental defectives, and brain-damaged adults—R. W. ZASLOW
An investigation of trait relationships among six-year-old children—S. A. AVAKIAN

VOLUME 64—July-December, 1961

1. Coping patterns of preschool children in response to intelligence test demands—A. MORIARTY
Studies of individual and paired interactional problem-solving behavior of rats: II. Solitary and social controls—A. BARON AND R. A. LITTMAN
2. Studies of role performance—J. H. MANN
A psychological assessment of professional actors and related professions—R. TAFT
A cross-cultural study of menstrual taboos—W. N. STEPHENS
The biological roots of creativity—H. GUTMAN

VOLUME 65—January-June, 1962

1. A Rorschach study of the development of personality structure in white and Negro children in a Southeastern community—A. C. PRICE
Measuring personality patterns of women—J. LOEVINGER
Home interests in early adolescence—M. AMATORA
Social and emotional adjustment during adolescence as related to the development of psychosomatic illness in adulthood—L. H. STEWART
2. A systematic study of interaction process scores, peer and self-assessments, personality and other variables—E. F. BORCATT
Children's conceptions of the content and functions of the human body—E. GELLERT

VOLUME 66—July-December, 1962

1. Ocular-manual laterality and perceptual rotation of literal symbols—I. FLESCHER
Parental child-rearing attitudes and their correlates in adolescent hostility—S. B. CHOROST
Some explorations into the nature of anxieties relating to illness—P. R. ROBBINS
Developmental-interactive aspects of hypnosis: Some illustrative cases—J. R. HILGARD AND E. R. HILGARD
2. Teaching a two-year-old to read: An experiment in early childhood learning—W. FOWLER
Achievement gains or losses during the academic year and over the summer vacation period: A study of trends in achievement by sex and grade level among students of average intelligence—K. M. PARSLEY, JR., AND M. POWELL

VOLUME 67—January-June, 1963

1. Verbatim dialogue of a mother and child in therapy—C. E. MOUSTAKAS
A study of self-concept: Delinquent boys' accuracy in selected self-evaluations—W. E. AMOS
An integrating study of the factor structure of adult attitude-interests—R. B. CATTELL AND J. HORN
Social factors in mental disorders: A contribution to the etiology of mental illness—W. G. COBLINER
2. Conceptual preferences associated with choice of quantitative and nonquantitative college majors—J. GILBERT
A study of the concept of the feminine role of 51 middle-class American families—A. STEINMANN

VOLUME 68—July-December, 1963

1. Brain dynamism as reflected in illusions—G. K. YACORZYNSKI
The nature and measurement of components of motivation—R. B. CATTELL, J. A. RADCLIFFE, AND A. B. SWENEY
2. Sharing in preschool children as a function of amount and type of reinforcement—W. F. FISCHER
The Gesell Incomplete Man Test as a measure of developmental status—L. B. AMES AND F. L. HIG
Psychological and sociological factors in prediction of career patterns of women—M. C. MULVEY
Functionalism and its historical significance—R. HARRISON

\$30.00 per year
\$15.00 per volume
Single Numbers \$7.50

BIMONTHLY
Three volumes per year

October, 1965
Volume 67, First Half

Founded in 1929 by John Dewey and Carl Murchison

THE JOURNAL OF SOCIAL PSYCHOLOGY

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

OCTOBER, 1965

(Manuscripts are printed in the order of final acceptance)

- Exploring eidetic imagery among the Kamba of central Kenya 3
By LEONARD W. DOOB
- Creativity, brainstorming, and orneriness: A cross-cultural study 23
By HENRY CLAY LINDGREN AND FREDRICA LINDGREN
- Quantitative aspects of the group-psychotherapist's role behavior: A method-
ological note 31
By JEROME RABOW
- A delay-of-reinforcement effect in persuasive communication 39
By RALPH L. ROSNOW
- The effect of "human" vs. "machine" set on group problem-solving procedures . 45
By JOHN K. HEMPHILL AND CAROLYN B. McCONVILLE
- Characteristics of nurses in short- and long-term treatment settings . . . 61
By JOAN K. BORUCHOW

(OVER)

Copyright © 1965, by The Journal Press
Provincetown, Massachusetts, U. S. A.

All rights reserved

Published at 35 New St., Worcester, Mass.

Second-class postage paid at Worcester, Mass.

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts

Some personality differences between dogmatic and nondogmatic groups . . .	67
BY WALTER T. PLANT, CHARLES W. TELFORD, AND JOSEPH A. THOMAS	
The definition and measurement of national morale and morality . . .	77
BY RAYMOND B. CATTELL AND RICHARD L. GORSUCH	
Social-desirability ratings of personality items by three subcultural groups .	97
BY CAROL JEAN DIERS	
The effects of discrepant information on perceptual processes: An application of balance theory	105
BY DAVID VERN STIMPSON	
Group responses to threat: Perception and learning under stress . . .	115
BY BRADFORD B. HUDSON, JOHN MCDAVID, AND MARIAN ROCO BINNER	
Some correlates of the don't-know response in opinion questionnaires . .	139
BY ROBERT C. ZILLER AND BARBARA H. LONG	
Perceptual accentuation as a problem in the psychology of judgment . . .	149
BY RAYMOND E. HARTLEY	
Role-deviant respondent sets and resulting bias, their detection and control in the survey interview	163
BY MERTON S. KRAUSE	
BOOKS	185
BOOKS RECENTLY RECEIVED	185

EXPLORING EIDETIC IMAGERY AMONG THE KAMBA OF CENTRAL KENYA*

African Studies, Yale University

LEONARD W. DOOB¹

A. INTRODUCTION

For many years it has been stated that eidetic images, defined in their visual form as images "persisting after stimulation, relatively accurate in detail, colored positively, and capable of being scanned" (6), are very rare among adults in Western society (e.g., 8, p. 651). Recently, however, the incidence of this type of imagery has been found to be significantly higher in a sample of Ibo in Eastern Nigeria (West Africa), especially among those living in a rural area (5). In the exploratory research to be reported, the author has sought tentative answers to five questions: (a) Does eidetic imagery occur equally frequently in another African society? i.e., what is its *incidence*? (b) If it does occur, what are its distinctive *attributes*? (c) Does it facilitate *accuracy* in reporting immediately after perception and thereafter? (d) With what psychological or social factors is it associated? i.e., what are its *relationships*? and (e) What functions does it perform from the viewpoint of people who possess it? i.e., what is its *ethnopsychology*?

Whenever possible, promising new leads were pursued, the most pressing of which suggested the possibility of another kind of imagery.

Since neither theory nor knowledge concerning eidetic images can indicate, alas, the kind of society in which they are likely to flourish, the investigator selected the Kamba for the fortuitous reason that he happened to be in East Africa for another purpose in the summer of 1964, and this tribe was close at hand. The Kamba (sometimes the plural form, Wakamba, is used to designate them) are a Bantu group numbering approximately 600,000 and living in

* Received in the Editorial Office on January 25, 1965, and given prior publication in accordance with our policy on cross-cultural research.

¹ This study is part of a research program on changing attitudes and values in African countries being supported by the Carnegie Corporation of New York. In addition, gratitude is expressed to Dr. and Mrs. Gordon Wilson of Nairobi for stimulating suggestions, friendly encouragement, and very practical assistance; to Mr. Nasson Mulili for functioning as an incredibly sagacious interpreter, gifted research assistant, and cheerful friend; to Mr. Juan T. Trippe for planning an imaginative project of another kind which, though unsuccessful, enabled me to travel again to East Africa; and to Professors Gordon W. Allport, Ralph N. Haber, and Elsa M. Siipola for penetrating ideas concerning the design of the study or for ruthless, constructive criticism of a first draft.

central Kenya. In culture and in language they are closely related to the better-known Kikuyu. The reservation on which the research was conducted is a flat, rather arid plain beginning about 20 miles northeast of Nairobi and having an altitude close to 6,000 feet above sea level. The Ss tested in the study live at least another 20 miles further east. The round adobe houses of the Kamba are fairly close together. The principal road is unpaved, reasonably good, usually very bumpy and, during the dry season, most unpleasantly dusty. Markets are held twice a week in small centers which have shops and other public services. There is no electricity in the area.

Three relevant characteristics of the Kamba language must be noted. Like many Bantu languages, it is deficient with respect to color terms; consequently, informants experienced some difficulty in naming colors and often used English terms that have entered their language. The distinction between present and imperfect tenses is as clear-cut as in English, although the two tenses may be used carelessly because they are often distinguished only by a single phoneme. Finally, the phrase, "picture in my head," which plays an important role in this study, can be translated literally from English; in fact the word for "picture" is generally one derived from the English term.

B. THE SAMPLE

Since a random sample of informants was neither necessary nor possible, the investigator and his interpreter simply approached a dwelling usually located not too far from the main road and asked the first adult they met whether he would be willing to look at some pictures and be questioned by "a professor from America." Men were generally sought because women either claimed to be occupied with household chores or were frequently distracted by young children; in some instances adolescents became the Ss. The very few people who refused seemed not suspicious nor unfriendly but in fact busy; some declined to be reinterviewed probably because their secret hope of being paid the first time had not been realized.

A total of 49 interviews was completed: 20 females, 29 males (six per cent more females than in the Ibo study); ages, not very reliably reported, ranged from 14 to 70, with a mean of 27.8 years (25.7 for the Ibo sample). Not all the findings, however, are based upon the total sample as the contents of the schedule were deliberately varied in the course of the investigation to adapt the procedure to the minor crises that arose and, more importantly, to take advantage of new information. For this reason, whenever the number of cases is *not* 49, the number giving rise to a statement is supplied in

parentheses, sometimes as part of a ratio indicating the proportion of *Ss* conforming to the indicated tendency.

C. FIELD CONDITIONS

With two exceptions, all interviewing and testing took place out-of-doors and, whenever possible, in the shade. The materials were placed upon a folding table, and *S* was provided with a folding chair. *S* sat directly across the table from the investigator, so that his behavior (especially his eye movements) could be closely observed. The interpreter was seated at one of the other sides of the table. Almost always, at least a half-dozen spectators of all ages stood around to watch and listen; infants, crowing chickens, placid dogs, and sometimes cattle also intruded. The general atmosphere, in short, was jolly, cordial, and incredibly noisy; most *Ss*, however, seemed able to concentrate upon the tasks at hand, probably because they were accustomed to living under these conditions.

Virtually all interviews were conducted in Kamba through the interpreter, but in a few instances *S* was addressed in English and replied in English. Verbatim notes were taken in English. The last eight interviews and all the follow-up interviews were also recorded on tape.

The field work progressed slowly. Living on the reservation was not practical; hence it was necessary to drive daily at least 90 miles to reach the site and return to Nairobi. A good deal of time was expended in procuring *Ss* and in responding to their endearing rites of courtesy and hospitality. The interviewing and testing of each *S* took at least an hour. Not more than four schedules were ever completed on a single day after nine hours of driving, searching, and testing. In the course of a month less than one-third of the 49 *Ss* could be located and tested a second time.

D. PROCEDURE

All *Ss* were tested for eidetic imagery in a standardized manner adopted from Haber and Haber (6) and also utilized in the Ibo study. After being given the first half of a visual retention test, *S* faced a blank piece of cardboard on a small stand. The classical procedure was then followed. To accustom *S* to seeing images on that screen, *S* was shown a red square for 10 seconds; then he was questioned about his after-image, which almost always appeared and turned out to be negative rather than positive. The training was repeated with a black square. Next, a drawing or photograph—hereafter referred to as *pictures*—was exposed for 30 seconds after the *S* had been

instructed to scan the entire stimulus. *S* reported immediately afterwards what he saw on the screen. If he could see nothing, he was asked to recall what he could from memory until he had supplied enough details either voluntarily or upon request to indicate with certainty that he had perceived the picture in the first place.

The first 24 *Ss* viewed a series of six pictures. At the outset these were four drawings and silhouettes from the Haber and Haber series and two photographs of definitely African scenes. After the seventh *S* this series was changed to include two pictures from Haber and Haber, two African scenes, and two complicated montages of various objects: animals, numbers, and letters.

At this point the procedure was drastically changed because most of the *Ss* had been reporting that ordinarily they did not see "pictures" in front of their eyes or themselves—rather they referred just to "pictures" without locating them or, more usually, to "pictures in my head." The remaining 25 *Ss* were first tested in the same way for eidetic images, except that the number of pictures in the series was reduced to four. The second half of Benton's test (3) followed, after which every *S* but one was shown a second series of four pictures and questioned as to whether or not he was seeing a "picture in your head." The two series were presented in alternate order from *S* to *S*; each series contained two pictures from Haber and Haber, one African scene, and one montage.

Since the traditional procedure eliciting eidetic images almost always makes use of a screen, those images elicited in the absence of the screen will hereafter be called pictorial images (PI) to distinguish them from eidetic images (EI). For reasons to be indicated subsequently, the pictorial images (PI) are considered to be different from though related to what the literature terms memory images (MI). An *S* seeing nothing on the screen or reporting no "picture" in his head will be said to have had no image (OI).

As in the Ibo study, each response to a picture was classified as E ("... correct minute details . . . [*S*] continued to scan the screen . . . as he reported that the image was fading or had faded, he changed tenses from the present to the past"); E- ("... the details were less numerous or the duration too short to permit a full report by [*S*] in the present tense"); or O ("... no trace . . . or a flat statement to the effect that 'I see nothing'"). A similar classification was adopted when *Ss* were asked to report whether or not they had picture imagery (PI), except that the criterion of duration could not be employed: P, P-, and O. Reliability was high—over 90 per cent—as determined in two ways. First, the investigator and the interpreter generally

whispered their judgments to each other as *S* was being exposed to the next picture. Then a sample of the recordings on tape was classified by two independent judges.

As in the Ibo study, each *S* was assigned to one of four EI groups and, for the *S*s viewing the two series of pictures, also to one of four PI groups. Criteria for assignment were slightly different from the Ibo study since not seven but either four or six pictures were used: (a) All E (or P)—no Os, no minus; (b) Mostly E (or P)—no Os and no more than one minus (if four) or two minus (if six); (c) Some E (or P)—no more than one or two Os (if four), one to four Os (if six); and (d) No E (or P)—three or four Os (if four), five or six Os (if six).

E. RESULTS

1. Incidence

The imagery groups in which *S*s were classified as well as comparable data from the Ibo research are given in Table 1. Virtually the same percentage of *S*s in both studies, 20 per cent, were found to be *All E*. The total having *All E*, *Mostly E*, or *Some E* among the Ibo was 53 per cent, whereas the

TABLE 1
CLASSIFICATION OF *S*s

Category	Ibo E	Kamba E	P
All E or P	9	10	12
Mostly E or P	6	2	7
Some E or P	9	7	3
No E or P	21	30	2
Total	45	49	24

figure among the Kamba was 39 per cent. The proportion of Kamba with PI was very high: almost all *S*s reported such images at least once (22 out of 24). As in the Ibo study, there is no statistically significant difference between the various pictures with respect to the frequency of EI elicited. One picture of an African scene, preferred by more *S*s than any other, produced 43 per cent EI defined as the sum of E and E— (16 out of 37); the corresponding figure for a non-African scene was 41 per cent (15 out of 36).

2. Attributes

*S*s reporting either EI or PI behaved differently from the way they did in the absence of such images:

1. EI and PI were described in the present tense; OI, in the past tense.

Occasionally an *S* with PI but never one with EI slipped into the past tense. When questioned concerning the shift, he immediately moved back into the present.

2. The eyes of an *S* with OI might wander in all directions while being questioned, but usually he looked directly at the interpreter or the investigator. In contrast, an *S* with EI concentrated upon the blank screen until the time he reported that the image had disappeared; and often spontaneously and always after prompting he would point to an approximately correct portion of that screen in order to locate one detail in relation to another. The modal behavior of *Ss* with PI seemed to be to avoid fixating upon a person and instead to stare blankly straight ahead and in some instances to close their eyes.

3. *Ss* with EI or PI were always willing and able to answer the question, "Is what you see on the screen [in your head] very clear?" The same question made no sense to *Ss* reporting OI. Not typically but from time to time, a few *Ss* stated that their images were "just as clear" as the original pictures. When an *S* could not supply the requested information, he might explain the deficiency in terms of the image he claimed to be experiencing. Concerning his PI, a young man said, "It is not appearing very well." A young woman commented on two PI she had been asked to observe simultaneously: "I see them but I cannot locate everything that is there." The sister of the *S* just mentioned stated that images derived from the same two pictures were "colliding."

Aside from their respective external and internal locations, EI differed from PI in two respects:

1. A majority of the *Ss* themselves distinguished between the kinds of imagery. They reported no EI during the first series of pictures, but then claimed to have only PI or at least some PI during the second series (15 out of 24). Seven of the remaining *Ss* had both EI and PI, and a pair had no imagery; but no *S* had EI without also having PI.

2. With no exceptions, the duration of PI was significantly greater than EI. Save for a single instance, all EI eventually disappeared. Eidetic images remained from 40 to 198 seconds, with a mean of 111 seconds ($N = 17$). The exception was a 15-year-old girl whose images, she stated, lasted as long as she wished to retain them. She was questioned for six minutes concerning one of the pictures that she claimed to be "seeing" as vividly at the end as at the outset. In contrast, with three exceptions, PI seemed never to disappear. *Ss* asserted that, after two to five minutes (when the investigator arbitrarily turned to the next picture or task), the images were just as clear or almost as clear as they had been immediately after exposure. In the case of one *S*,

questions about the PI were posed for eight minutes—in the case of another for 12 minutes—and still, the *S* maintained that the images persisted. Only one of the three exceptions stated that all PI disappeared (mean = 118 seconds). The other two *Ss* reported that, except for one picture each, their other PI were either persisting or becoming unclear.

3. Accuracy

Although individual *Ss* with EI or PI displayed impressive feats of memory, which anecdotally might be mistaken for minor miracles, careful and varied analyses reveal that *type* of imagery had virtually no effect upon the accuracy of report in the absence of the pictures. Recall was measured at three intervals:

a. Immediate recall. After seeing each picture, as indicated earlier, *S* was questioned about his EI, PI, or OI.

b. Short-time recall. After being given (during the initial interview) a series of tests and questions that lasted from 14 to 66 minutes, *Ss* were again asked to look at the blank screen and report whether or not they saw any of the pictures thereon ($N = 46$). Next they named the picture they liked best. Then they were told to recall whatever pictures they could, either on the screen or as pictures in their heads; and finally, after spontaneous recall had been exhausted, eight *Ss* were asked whether or not "you can see on the screen or as a picture in your head the one in which . . ."—and there followed a brief description of the picture, roughly in the words used by the *S* immediately after exposure.

c. Long-time recall. From six days to over three weeks after the first interview, with the mean elapsed time being slightly under two weeks, 17 *Ss* were located and interviewed a second time. First, they were given an irrelevant icebreaker concerning daydreaming. Then they were shown the same blank screen and asked to report whether or not "you can see on it any of the pictures we showed you when we were here last time." In the absence of EI (or after the appearance of the last EI), they were questioned concerning possible PI; next they tried to reply to the simple question, "which ones can you remember?" These invitations to spontaneous recall were followed by prompting inquiries about the remaining pictures.

Two modes of comparison and various measures of recall were employed in connection with the three types of recall. For immediate and short-time recall, the reference point was *S*: nineteen *Ss* with EI (i.e., the sum of *All E*, *Mostly E*, and *Some E*) *vs.* 30 with no EI and seven *Ss* with both EI and PI *vs.* 15 *Ss* with PI but with no EI. The *Ns* were smaller for short-time recall because the section of the interview containing this test was omitted in

some instances. Then recall data were also broken down in terms of the three categories of images, and a picture-by-picture analysis was made.

Measures of recall were those that took into account the fact that the amount of recall was influenced by the number of questions that were asked. Up to a point, additional information could almost always be obtained by further inquiries; also the number of questions varied, inasmuch as Ss provided information with varying degrees of agility. Two of the measures were ratios. One ratio shows the relation between the number of content items (or colors) correctly reported and the number of appropriate questions posed by the investigator. In addition, two specific details were singled out for each picture (e.g., the number of branches on a tree, the number of men in a group having hats on their heads), and note was made of the correctness of the information supplied either spontaneously or after prompting. Appropriateness and feasibility (especially the number of cases available) determined which of the measures was used to measure a given type of recall. For short-time and long-time recall the number of pictures recalled spontaneously or after prompting was used as another measure.

Although the findings largely were negative, certain observations and modal tendencies deserve brief mention. For *immediate* recall the mean number of correct items recalled per inquiry was calculated for the three kinds of imagery evoked by each picture. The absolute differences between those means were extremely small; hence statistically are nonsignificant, yet (without exception) the mean for OI was always greater than the one for either EI or PI ($p < .01$, sign test). In addition, fewer Ss with EI than those with OI spontaneously referred to specific details concerning the pictures ($p < .01$, sign test); but this trend does not appear when comparisons are made between those with PI and OI or between Ss with either EI or PI and Ss with OI under conditions of prompting or of the combination of spontaneity and prompting. Most Ss correctly recalled colors, although those not acquainted with English color terms often had difficulty expressing themselves; hence pointed to the color of an object at hand in order to designate what they meant. Almost no negative afterimages or memories were reported.

The ability of Ss to recall minute information was most easily tested by means of four pictures containing numbers or letters. Obviously only those claiming sufficient literacy were questioned. The two cars in a crowded African scene had license plates with letters and numbers clearly visible. Among eleven Ss with EI, only one recalled both series perfectly; two, almost so. The corresponding figures for five Ss with PI were two and one; and for nine Ss with OI, zero and one. Included among five objects in a

montage were 11 random numbers preceded and followed by zeros. Nobody recalled these numbers without some error (usually a gross one). The best performance among five Ss with EI was recalling correctly the first five numbers; among seven Ss with PI, the first three numbers and a pair in the middle; and among 10 Ss with OI, the first three and the last two numbers. Another montage consisted (a) of four random letters at the top and another four at the bottom of the various stimuli and (b) of four numbers under the figure of a buffalo. Nobody in any of the three imagery groups could correctly recall both sets of letters and numbers ($N = 21$), but a sprinkling in each group reported only one of the series correctly. Finally, a dozen random numbers ($N = 6$) or colors ($N = 4$ illiterates) were exposed to Ss with PIs. Nobody could repeat correctly all or even most of the numbers or colors either backwards or forwards.

The accuracy of *short-time* recall was unrelated not only to the imagery in immediate recall but also both to the imagery during that recall and to the order in which the pictures originally had been presented. The unimportance of the latter factor is shown by the Ss' favorite picture (a scene in an African town: two cars, a large bus, many people) that spontaneously was recalled equally well whether it had appeared in the first or the second series during exposure. Almost all Ss recalled spontaneously at least one picture (43 out of 46). The recall, whether spontaneous or prompted, took various forms. Ten Ss reported EI at least once, eleven reported PI at least once, and the remainder exclusively OI. Five Ss had EI of pictures which had *not* evoked such images immediately after seeing them: three previously had claimed PI; and two, OI.

During the second interview when *long-time* recall was tested, every S recalled at least one picture spontaneously with some degree of accuracy ($N = 17$). For one, more, or all of the pictures, seven of the Ss reported that recall was based upon scanning EI. Four of them previously had indicated they were seeing EI immediately after perception, and two and one had indicated, respectively, PI and OI.

Long-time recall was associated, not unexpectedly, with a factor unrelated to imagery: viz., reinforcement during the original interview. Whenever a preference for a picture had been expressed during that interview, the picture was spontaneously recalled in the second interview (12 out of 12). Almost without exception, Ss who had recalled pictures later in the first interview also recalled them in the second one (12 out of 14); yet the same proportion of the same Ss also recalled one or more pictures that they had not mentioned as part of their short-time recall in the first interview.

Identical tendencies appeared with reference to details recalled from the pictures. Most but not all the material previously had been reported immediately after perception or as part of short-time recall. Of the four pictures containing numbers or letters, only one evoked more than fragmentary recall: the African scene containing two cars with license plates. Four *Ss* were involved: one with PI after perception recalled both plates correctly, one with EI recalled only one plate correctly; yet another with EI recalled only a fragment of one plate, and one with OI reported fragments from both plates. All four *Ss* recalled, whenever they did recall something, in precisely the same way as they had reacted previously immediately after seeing the pictures, except that in the case of one *S* with EI the second license combination had dropped out. In contrast, the *S* who had reported OI immediately after exposure and who had not been questioned concerning possible PI recalled as PI in the later interview a number of items he had *not* originally noted. He stated that the mouth of the fish was "closed" without having made any previous reference to that mouth. He named the correct number of branches on a tree when, previously, he stated he could not remember the number. Another *S* had incorrectly reported a girl's hair to be "reddish" immediately after viewing it. Twelve days later, he noted correctly in his PI that it was "yellow."

4. Relationships

Relationships between *Ss* with and without PIs could not be analyzed because too few fell into the latter category; so three series of comparison have been made between those having *All E* or *Mostly E* and those having *Some E* or *No E*:

a. *Performance on various tests.* Six tests revealed no significant differences. They are as follows.

(1). *Sorting test.* A simple sorting test in which *S* is instructed to place 16 pieces of cardboard (differing with respect to shape, size, color, and insignia) into as many piles as he wishes, so that the pieces in each pile have "something in common" (4, pp. 316-17). There was no difference in the accuracy or the type of performance. *Ss* with EI did not have a tendency to use color alone or in combination with some other attribute as the organizing principle ($N = 46$).

(2). *Benton's Test of Visual Retention, Multiple-Choice Form (3).* *S* looks at a geometrical figure for 10 seconds and then, after its removal,

tries to find it among four similar drawings, only one of which is the correct one ($N = 37$).

(3). *An olfactory test.* At random S is offered one spice to smell and is then told to identify it among three similar spices all of which are presented in identical containers. Two trials are given ($N = 14$). Later in the interview, 10 S s were asked again to locate the same two scents. The tests were uniformly too easy during the initial presentation and too difficult for later recall.

(4). *Repetition test.* The ability to repeat immediately afterwards ($N = 12$) or 10 minutes later ($N = 7$) some simple, short sentences in German, a language unknown to the S s.

(5). *Color-memory test.* Two series of eight colored crayons were alternately exposed for 15 seconds, they were shuffled, and S was asked to restore them in the same sequence in which they had been shown him ($N = 14$). One series contained colors that are easy and the other those difficult to name in the Kamba language.

(6). *Thematic-apperception test.* The quantity and quality of responses to a verbal TAT ($N = 11$) or to TAT-type instructions concerning two small rubber figures ($N = 5$).

b. Census-type information. None of the following were related to the possession or nonpossession of EI: age (some S s, like Africans in many societies, were unsure of or could not supply this datum concerning themselves); sex; religion; age at which S s, if literate, believed (or guessed) they had been taught to read and write; unreliable, stereotyped information supplied by the S s themselves concerning their contact with the mass media of Kenya; length of time spent in the city of Nairobi, the very metropolitan capital of Kenya. Of the nine S s who stated they had never been in Nairobi two had strong EI; two, weak EI; and five, none at all. On the other hand, when those S s having any kind of EI are compared with those having no EI, one finds an almost significantly larger number of the former to know no English or virtually none and to have had either no schooling or less than three years ($.10 > p > .05$, χ^2 , two-tail test).

c. Claims concerning other kinds of imagery. No differences existed with reference to each of the following:

(1). *Modalities appearing in dreams.* Most S s claimed to experience colors (33 out of 37), sounds (24 out of 30), smells (21 out of 35), and tastes (24 out of 29) in their dreams.

(2). *Material gained through other modalities that might be stored as PI.* Most Ss stated that they had such images concerning sounds (32 out of 38), smells (27 out of 36), and tastes (21 out of 36).

(3). *The part of the body in which S believed his "self" to be* ($N = 22$). Most Ss mentioned the whole body or "every part of me" ($N = 10$); brain, mind, head ($N = 9$); and heart ($N = 8$).

5. *Ethnopsychology*

Among Ss with EI, no one claimed ever to have had the experience of seeing images on a blank surface (such as the testing screen) or just in front of their eyes. In fact, the kind of surface used is not available to Ss in a traditional village because they live in huts having walls relatively dark and certainly irregular. Projecting images on to a screen seemed "unnatural." Most of those questioned about the fact that they had reported EI upon the blank screen said they were surprised to see the "pictures" there (10 out of 12). Unless otherwise indicated, the ethnopsychological material presented in this section refers only to PI, which the investigator usually called "pictures" while interviewing the first 24 Ss and always called "pictures in your head" thereafter and in all the reinterviews.

All Ss (including those reporting only OI and not tested for PI), when asked directly about EI, pictures, or PI, answered questions freely and easily about the nature of the "pictures" ($N = 43$). Most maintained that "generally," "ordinarily," or "frequently" they experienced PI (25 out of 29). In fact, among the few Ss who could be questioned, all claimed to have seen one or more of the investigator's drawings or photographs between the first and second interviews ($N = 9$). On eleven occasions, at the very end of the interview, the villagers who had been watching the S were asked en masse or one by one whether they themselves also had PI. With only one exception, the majority of those present laid claim to having the images. They replied as seriously as they did concerning any other factual matter. Possibly the number of cases having some form of PI might be increased from 22 (see Table 1) to about 56, and those not having them from two to about six.

Every S except one ($N = 38$) unhesitatingly stated that PI were useful. Most stressed some general function, such as helping one remember events or people's faces from the past ($N = 16$); seven believed the pictures made them feel happy, and five believed they helped with school work. "I know what it looks like when I am here," an S said of a street in Nairobi which he claimed he was viewing in his PI. Some Ss stated that PI possessed movement: for example, in recalling a traditional Kamba dance (four out of six). The rest claimed that PIs were always stationary. Almost all Ss reported, when

queried directly, that they saw such pictures at night just before going to sleep (33 out of 45). A handful, while being interviewed concerning PI in general, slipped easily and apparently unwittingly into a discussion of their dreams. A small number stated that they could insert themselves into any PI (seven out of 13): for example, each of two men (when provoked by questions) could "see" himself besides the figure of the man fishing from the canoe in his PI of one of the pictures. It is perhaps significant, since traditionally the Kamba attributed great power to ancestral spirits (10, pp. 177-205), that the majority stated, when questioned, that PI enabled them—for years and years, some maintained—to recall dead persons as they had been at the time of death (19 out of 23); but six *Ss* could also recall dead people as they had looked earlier in life. After one *S* announced that he knew the investigator would be returning to interview him that day because a PI had so informed him, other *Ss* were asked whether PI also helped them anticipate the future. A majority (six out of 10) were convinced that PI performed that function and mentioned specific instances; but the rest had no faith or only partial faith in them. *Ss* tended to plead ignorance concerning whether or not "other people have pictures in their heads the way you do" (eight out of 10).

The majority of *Ss* stated that sometimes PI appeared spontaneously and at other times were voluntarily evoked (eight out of 15). The remainder claimed either mode but not both of them. In fact, all *Ss* indicating EI or PI after the original perception and a few claiming only OI, reported (either spontaneously or after prompting) the reappearance or the first appearance of at least one EI or PI as part of short-time or long-time recall. In this sense, the arousal of the images was "voluntary," although of course it was cued by the investigator's instructions. The images spontaneously recalled after the instruction, "Which of the pictures we showed you can you now see on the screen [in your head]?" were reported to be present almost instantly. In a few cases, however, *Ss* paused for 20 seconds or so before providing a verbal description. "Let me think," one *S* said concerning a picture that originally had produced an EI 13 days before; then he became silent while apparently concentrating; finally he emerged with a PI.

It is possible that *Ss* stressed the importance of PI since they knew of the investigator's interest in the subject. With an open-ended question concerning how *S* remembered a friend in his absence ($N = 41$), only a small number mentioned PI voluntarily ($N = 9$); but a large group ($N = 29$) also referred to those images when asked directly whether or not they utilized PI for that purpose.

One formal attempt and a number of informal attempts were made to check the validity of the belief in the utility of EI and PI. During the interview itself, as already indicated, the Benton Visual Retention Test was split into two equivalent parts, one of which was administered immediately before and the other immediately after the series of pictures testing for EI. The order of presenting the two halves was alternated from *S* to *S* ($N = 37$). Would the arousal of EI by the pictures carry over to the Benton task? In fact, slightly more of those with EI than of those without EI improved after the first picture series (*All E* and *Most E* vs. *Some E* and *No E*), but that difference is not significant ($p < .10$, one-tail test). With six eidetikers, the blank screen was returned during the second part of the Benton test, and *Ss* were instructed to retain the EI thereon while picking out (from among the four drawings) "the one you have just seen." This procedure led to no significant improvement.

Rough empirical tests on a few *Ss* also failed to substantiate claims concerning the utility of the EI. Two *Ss* who said they were guided by EI in free-style drawing were asked to reproduce the pictures giving rise to the EI they were experiencing. Both drawings were miserable failures; nevertheless the *Ss* insisted they were being guided by PI. Another *S* whose PI was remarkably detailed and accurate claimed that she could "see" in her PI whole pages from books. When tested by being asked to recall (a) pages she had originally read in school weeks before and (b) a new page that she had viewed for nine minutes in the presence of the investigator, she committed gross errors and showed evidence of employing ordinary memory (e.g., she substituted synonyms for words in the original text of the new page, and she paraphrased or omitted unimportant words). On the other hand, when six *Ss* were asked to reinstate some scene from the past (such as a main street in Nairobi), they claimed to be able to do so and, when questioned, immediately began reporting details with a high degree of accuracy and in the present tense.

F. DISCUSSION

This study indicates that a noteworthy proportion of *Ss* in a haphazard sample of adolescents and adults among the Kamba possess EI according to the usual definition of that imagery. The incidence is roughly the same as that in an Ibo sample and is much higher than that reported among European adults. Once again the reality of the phenomenon has been demonstrated, and the usual criteria—vividness, clarity, duration, positive afterimages, projection, scanning, the use of the present tense in reporting details—have been found both relevant and satisfactory.

The much higher incidence of PI among these Kamba, however, raises the challenging question of the relation of such imagery to EI and MI. As indicated, PI differed from EI with respect to subjective location and reported duration; but they were similar with respect to the tense in which they were described and also in the vividness ascribed to them. Without exception, all Ss claiming PI maintained they were seeing "the picture in my head" then and now; when questioned, they believed that having a PI of a picture was different from remembering it ($N = 14$). Thus some Ss were shown successively two montages not in the regular series: each displayed eight miscellaneous objects, some of which were similar but not identical (e.g., in one picture the color of a sock was orange and red; in the other, white) and each was distinctively labelled by referring to one drawing in the center (an ostrich and a giraffe: two familiar animals). One S who was interviewed in English could unerringly indicate, during the period of short-time recall, the picture in which a designated object had been displayed by consulting, he said, PI both of which he simultaneously brought back when instructed to do so; but one detail was missing:

What about the spade, can you see the spade?

I didn't see the spade.

You haven't seen the spade?

Not yet.

Well, how does it happen that in both the ostrich picture and the giraffe picture you have seen everything except the spade?

Yes.

Why was that, do you think, why, why didn't you notice the spade, why can't you see it now?

Because I didn't see it.

You didn't see it—well, then, you see in the pictures in your head only what you have seen—

No, I see them now.

You see them now, you see them now, but you don't see the spade.

I didn't see it.

The tape must be heard to appreciate the vehemence with which this young man defended, in the italicized sentence, the subjective reality of his PI.

No hard-and-fast conclusion can be drawn about the PI, which clearly have elements in common with both EI and MI. Even Jaensch (7, p. 16) has said that generally EI "merely have the character of *pictures*, which, nevertheless, are seen in the literal optical sense." The possibility that PI are nothing more than MI cannot be excluded. The Kamba Ss belong to a different culture; hence may have been encoding their internal responses in a manner that misleadingly suggested resemblances to EI. Tentatively, the

author thinks of PI as a form of MI that retains almost all the usual attributes of EI or as a form of EI that is coming to resemble MI. Such a guess springs from the promising but unproved assumption that EIs are modally present in traditional Kamba society and from the fact that MIs are the modal images in Western society; consequently, at this historical moment the Kamba may be in a transitional stage. Even the completely illiterate Ss in the sample who had never visited Nairobi nevertheless had frequent contacts with many aspects of Western civilization. They constantly saw automobiles, they heard loud-speakers blare in their markets, they were influenced by Roman Catholic and Salvation Army missionaries active in their area, they received orders and advice from government officials who exemplified some Western customs, a few had radio sets at home, etc.

Careful questioning and various types of analyses have revealed that the accuracy of recall was by and large unrelated to imagery. This finding is surprising. Ss in published research (e.g., 2, p. 424; 9, p. 192), as well as many Ss in the present study, report that EI or PI help them store information. Indeed, again and again other investigators have described Ss who are able to use their images to read materials (such as words from a foreign language) which presumably they could not have memorized while perceiving the original stimulus (1, p. 109); yet a definitive summary has suggested:

It happens only exceptionally that all the details of the stimulus object are correctly reported in the EI. There seems to be a continuum from a "blurred shadow" to almost photographic fidelity (8, p. 658).

All that can be said, in commenting upon these conflicting findings, is that Kamba Ss (when failing to recall certain details either from EI or PI) without exception replied in one of two ways. At some time during the testing all expressed themselves in the present tense: "I see the numbers but I don't know what they are," "I do not see it now," "It is not clear," "If I remember something, I see it." Or else they would say in the past tense that "I forgot to see" certain items during the original exposure. Typical is the following inquiry concerning a PI:

Do you see this picture in your head?

Yes.

Is it clear?

Yes.

What can you see?

I see three cars and people.

Can you see the license numbers on the cars?

I see them, but I did not observe them.

Can you observe them now?

Yes.

Can you tell us what the numbers are? Look at the picture in your head carefully and tell us what the numbers are, please.

I did not look carefully at the numbers.

Can you look and see them now?

No.

Another *S* complained, when he could not report the letters and numbers from a montage he had just perceived: "I wanted to put them in my mind, but then you took the picture away so quickly." The *S* above who could not see the spade in his PI of the "ostrich" and "giraffe" pictures was then shown the original pictures a second time and the spade was pointed out to him; thereafter, he indicated and demonstrated that both of his PI contained that object. In the present study there were no instances, as reported by other investigators (8, p. 648), of the gradual appearance in EI of details which at the outset *Ss* claimed to be unable to "see."

In brief, the contents of EI or PI seemed to include only information that had been perceived as carefully as any subsequently stored through some other type of encoding. In addition, that storage resembled other forms of memory. By and large, it has been shown that what emerged in long-term recall was repetition of the immediate report and of short-time recall. *Ss* with such images tended to group numbers and frequently reported letters not in their original order. Nobody could "see" very efficiently either groups of letters or numbers, especially when instructed to report them backwards. Fragmentary evidence indicating the later recall of information not originally reported or recalled and a shift in imagery over time, however, suggests some of the complications that may eventually need to be analyzed in studying retention.

The statistically significant tendencies during immediate recall for *Ss* with OI to provide more information per inquiry than those with EI or PI and for *Ss* without EI to provide more information spontaneously than those with EI suggests the possibility that EI is a *less* efficient way of storing or coding information. Thus, some *Ss* with OI gave long and detailed reports concerning what they had seen; whereas generally *Ss* with EI or PI required a series of questions to elicit the same amount of information. It looks as though the ability to verbalize may have been playing an important role, and such an ability is likely to be positively correlated with schooling which in turn has been found to be negatively (if not significantly) correlated with

the incidence of EI. At any rate, perhaps Ss with OI had already coded their impressions in a readily reproducible, coherent form. In contrast Ss with EI or PI could continue to "see" details in the images facing them or inside their heads while for the first time, in whole or in part, they transformed those images into words in order to respond to the investigator's questions.

In addition, the images of the Kamba may not have been more efficacious since they were being employed to perform an unusual task. Were this so, then accuracy ought to have been greater in recalling details from the more familiar African photographs than from the Western-style drawings of Haber and Haber; but no differences emerged. If before the coming of the British and other Europeans Kamba imagery functioned as it does now, then hypotheses suggesting that EI help nonliterate people store essential information about themselves and their society must be tentatively rejected. The present contacts of the Kamba with Western society, however, may be the cause of a decline in the use of imagery to mediate recall. Or else the Kamba, both formerly and now, may be like Western Ss. EIs are strong among children and not among the adolescents and adults composing the sample that was examined.

That Ss with EI could not be differentiated from those with OI on a number of tests and inquiries is contrary to Schachtel's and Werner's speculative views (*see* 11, p. 160; and 12, pp. 145-146). These investigators believe that EIs are part of a syndrome reflecting a mode of making basic perceptual adjustments to the environment. Of course such an hypothesis might eventually be vindicated by tests other than those used in the present research. Likewise, the largely negative findings reported here suggest the need for further investigation of the relation between EI and ecological and social factors (unquestionably with larger numbers of Ss). The almost significant differences that were obtained point in the same direction as do the Ibo results. Contact with some aspect of the West (a town for the Ibo, English or education for the Kamba) apparently decreased the incidence of the imagery.

The data on the ethnopsychology of PI clearly show that the imagery was meaningful to people. What many Kamba Ss said about the pictures in their heads suggests that the PI may perform mediating functions other than promoting the storing of information. Details recalled in the present tense (immediately or weeks after the original exposure) have perhaps more "warmth" or "meaning" or they seem more "real" or "compelling"—whatever it is that those terms mean phenomenologically—than ones described as having been viewed in the past. Is the PI of a deceased ancestor, though not in the least hallucinatory, more realistic than an MI thereof? PI, possibly also EI, may have once provided information to people lacking visual media (such

as paintings or print) and thereby have enabled them to preserve their impressions from the past not more accurately but more vividly; and they still may be serving that function. Indeed, the author has the undocumented impression that Ss with EI or PI answered questions about the pictures they had viewed with greater certainty than those lacking them. As would follow from their ethnopsychology, people who believe in the efficacy of the images are likely to feel confident while experiencing them. Such hypotheses must be cautiously advanced, however, in view of the fact that in this study *type* of imagery was not associated with expressed preference for pictures. Of the 14 Ss who liked an African scene best of all, five reported it had produced an EI; two, a PI; and seven, OI.

Possibly the most important problem to be investigated in future research is the isolation and identification of the factors in the milieu and in the individual's own life experience that reinforce or extinguish this ability to have or to believe one has EI or PI. Even if the negative correlation between the possession of EI and contact with Western culture that was noted clearly in the Ibo study and less clearly among the Kamba were confirmed, it would still be necessary to locate the precise experiences that encourage the images in traditional African societies and not in Western societies. For those Africans who become acculturated during their own lifetime and yet continue to report EI or PI, the conditions in early childhood that make the ability so strong that it can withstand in the future the forces that ordinarily destroy them must be established. Do those conditions involve the age at which contact with the West (e.g., learning to read or to encode experience in new ways) has taken place? Are EI and PI replaced by MI when individuals are less dependent upon and feel themselves better able to cope with the external world?

The process of learning and unlearning the proclivity toward EI or PI may very well differ from the usual kind of social learning in which parents or peers deliberately or unconsciously employ a system of direct rewards and punishments to affect or control *observable* behavior. In contrast, images are obviously private, *nonobservable* phenomena that can be only indirectly affected through external events. A growing boy can be shown how to hold his hand in order to discharge a weapon, but he can scarcely be given (even if someone wished to do so) comparable tutelage concerning the pictures in his head. On a purely anthropological level, it will be important to investigate whether the existence or incidence of EI and PI is correlated positively with some customs and beliefs (e.g., ancestor worship) or negatively with others (e.g., a system of mnemonic devices for recording events).

G. SUMMARY

On the basis of conventional and rigorous criteria, 20 per cent of a sample of Kamba adolescents and adults in Kenya ($N = 49$) were found to possess eidetic imagery. That proportion, also emerging in another African society (the Ibo of Nigeria), is many times higher than the incidence in the West. Almost all the Kamba so tested claimed to have "pictures" in their heads of past events: a kind of imagery shown here to be both different from but related to eidetic and memory images. Neither eidetic nor pictorial imagery aided immediate, short-time, or long-time recall. If anything, the absence of these images was associated with better recall right after seeing the stimulus pictures. The presence of eidetic imagery was not associated with performance on a number of psychological tests nor with most census-type information, although there was a slight tendency for it to be negatively correlated with knowledge of English and formal schooling. Most Ss claimed that pictorial images are helpful to them in their daily lives: for example, by enabling them to recall the dead. The tentative conclusion reached is that eidetic and pictorial images may enable people to recall the past not more accurately but more vividly and confidently.

REFERENCES

1. ALLPORT, G. W. Eidetic imagery. *Brit. J. Psychol.*, 1924, **15**, 99-120.
2. ———. The eidetic image and the after-image. *Amer. J. Psychol.*, 1928, **40**, 418-425.
3. BENTON, A. L. A multiple choice type of the visual retention test. *A. M. A. Arch. Neurol. & Psychiat.*, Chicago, 1950, **64**, 699-707.
4. DOOB, L. W. *Becoming More Civilized*. New Haven: Yale, 1960.
5. ———. Eidetic images among the Ibo. *Ethnology*, 1964, **3**, 357-363.
6. HABER, R. N., & HABER, R. B. Eidetic imagery I: Frequency. *Percept. & Mot. Skills*, 1964, **19**, 131-138.
7. JAENSCH, E. R. *Eidetic Imagery and Typological Methods of Investigation*. New York: Harcourt, 1930.
8. KLÜVER, H. The eidetic child. In C. Murchison (Ed.), *A Handbook of Child Psychology*. Worcester, Mass.: Clark, 1931. Pp. 643-668.
9. ———. Eidetic phenomena. *Psychol. Bull.*, 1932, **29**, 181-203.
10. LINDBLOM, G. *The Akamba in British East Africa*. Uppsala: Appelberg, 1916.
11. SCHACHTEL, E. G. *Metamorphosis*. New York: Basic Books, 1959.
12. WERNER, H. *Comparative Psychology of Mental Development*. New York: International Universities Press, 1957.

Department of Psychology
Yale University
New Haven, Connecticut

CREATIVITY, BRAINSTORMING, AND ORNERINESS: A CROSS-CULTURAL STUDY*

Department of Psychology, San Francisco State College

HENRY CLAY LINDGREN AND FREDRICA LINDGREN

A. INTRODUCTION

In recent years the topic of creativity has been attracting an increasing amount of attention on the part of leaders in the fields of government, industry, and education. There seems to be a widespread feeling that yesterday's methods are not applicable to today's problems, that new imaginative solutions are needed—solutions characterized by ingenuity and, of course, by creativity. This urge to seek creative solutions appears to be especially acute among those who work with and in the developing countries.

An important question is, "can creativity be developed, stimulated, or otherwise encouraged and facilitated or are we dependent on those genuinely creative individuals, who apparently are quite rare and unable to stretch their resources to meet the demand for creative problem solving?"

One of the more prominent proponents of the idea that creativity can be stimulated is Alex F. Osborn (5), who in 1939 introduced a technique he called "brainstorming," an approach to problem solving that makes use of uninhibited group discussion. With brainstorming, all ideas put forth by a group member are accepted without question, and no ideas can be criticized; in fact, members are urged to build on one another's ideas. Osborn maintains that brainstorming has a stimulating and facilitating effect and that people can produce twice as many ideas in such groups than they can working alone.

Although at least two research studies (2, 6) have shown that individuals brainstorming alone are more creative than when brainstorming in a group, data from one investigation (2) imply that the creativity expressed when brainstorming alone can be increased *if preceded by group brainstorming*. This finding was confirmed by a study (4) conducted by the present authors, with groups of students at San Francisco State College. The authors' study consisted of a three-phase experiment in which subjects were asked to write captions for cartoons alone; then, in a group session held under brainstorming rules; finally, alone. The level of creativity of the captions written during the

* Received in the Editorial Office on May 3, 1965, and given prior publication in accordance with our policy on cross-cultural research.

third phase (after the brainstorming) was judged to be higher than that of those written during the first phase. The number of captions produced during the third phase was also higher, although the increase was statistically significant only for women.

In the foregoing study, the authors used a forced-choice questionnaire (the Asymmetrical Preference Test or APT) based on two sets of self-descriptive adjectives: one set chosen by persons preferring asymmetrical art; the other, by those preferring more-symmetrical or conventional pictures (1). In general, subjects preferring asymmetrical art characterized themselves as asocial—that is, as “ornery,” “cantankerous,” and “difficult”—whereas those preferring symmetrical art described themselves as peaceful, agreeable, and responsive to the needs and feelings of others. When the APT was scored according to a key based on choices made by persons preferring asymmetrical art, the resulting scores correlated positively and significantly with the level of creativity of captions produced during both the first and the third phases of the experiment. Correlations of APT scores with the number of captions produced were not significant. The authors concluded that in spite of the fact that the self-regarding attitudes of the more-creative subjects had a pronounced asocial bias, these individuals nonetheless were able to benefit from the socially stimulating experience of brainstorming.

The purpose of the present study was twofold: (a) to determine whether or not brainstorming would stimulate creativity among members of Middle East cultures, who ordinarily function in non-English languages and (b) to determine whether or not the kinds of self-regarding attitudes characteristic of more-creative American subjects (and measured by the APT) characterize subjects from a Middle East culture.

The design of the original study had not included a control group, thus leaving unanswered the question as to whether or not gains in level of creativity could result from practice rather than from group brainstorming. The present study included a control group in order to answer that question.

B. PROCEDURE

Subjects participating in the study were 77 men and 28 women in nine psychology classes at the American University of Beirut, Beirut, Lebanon. Two additional classes made up a control group consisting of 29 subjects (19 men and 10 women). Most of the subjects came from the various Middle East Arab countries (Lebanon, Syria, Kuwait, Iraq, Saudi Arabia, Jordan, and Sudan), with a few from Turkey, Cyprus, Iran, and Pakistan.

All class members present on the day of the experiment were permitted to participate, but the response sheets of all those listing English as their first language were eliminated from further consideration. The statistics reported are based solely on the responses of subjects who reported some language other than English as their primary tongue. Arabic was the language more commonly reported as a first language; Armenian was second.

In Phase I (individual brainstorming) of the experiment, subjects were given a free-form cartoon and were asked to write as many humorous or interesting captions (or "whatever comes to mind") as they could in a six-minute period. In Phase II (group brainstorming), subjects were organized into groups of four or five, given a second cartoon, and told to follow the same instructions as for Phase I, except that they were to function for 10 minutes as a discussion group operating under brainstorming conditions. In Phase III (individual brainstorming) subjects were given a third cartoon and told to write captions, following the same instructions and observing the same time limit as in Phase I of the experiment.

After the completion of Phase III, subjects were asked to complete the APT. They were given Arabic translations for some of the more complicated terms, and further definition was provided orally on request. In spite of the fact that the questionnaire was not in the native language of the respondents, it proved to be almost as reliable as it was when it was used with San Francisco subjects in the original study (.86 *vs.* .89).

Three judges operating independently rated each of the response protocols for Phases I and III for overall creativity. Judges were instructed to base their ratings on the extent to which protocols could be characterized as being original, appropriate, humorous, or interesting. Protocols were first sorted randomly into batches of 16, and each batch was rated on a seven-point scale in such a way that judgments were forced into an approximation of the normal curve. One response sheet in each group of 16 was assigned the rating of "1" (least creative), and one was assigned a rating of "7" (most creative); two sheets were assigned the rating of "2," and two were assigned a rating of "6"; and so on. When each judge had completed his rating of all protocols, the protocols were randomly sorted before new batches of 16 were prepared for the next judge. This resorting of protocols probably reduced the inter-judge reliability because each protocol was compared with up to 15 new protocols each time it was judged. The resorting, however, did help compensate for the fact that a mediocre protocol might be overrated because it happened to appear in a generally uncreative batch or might be underrated

because it appeared in a generally creative batch. In spite of the deliberate introduction of such random elements, the interjudge reliability was high. Kendall's coefficient of concordance is a highly significant .76. Judges' ratings were summed to give each subject two level-of-creativity scores: one for his Phase I protocol; one for his Phase III protocol.

C. RESULTS

Table 1 presents results that in the main are consistent with those obtained with the San Francisco sample on which the original study was based. The experimental group made significant gains both in the number of responses and in the level of creativity, whereas the control group did not. Indeed,

TABLE 1
DIFFERENCES IN THE NUMBER OF RESPONSES, LEVEL-OF-CREATIVITY SCORES, AND
ASYMMETRICAL PREFERENCE TEST (APT) SCORES

Variable	N	Phase I		Phase II		t
		Mean	SD	Mean	SD	
Number of responses						
Men	77	4.36	2.41	5.07	2.41	2.84*
Women	28	4.61	2.05	6.17	2.19	3.31*
Control	29	5.12	2.18	5.22	2.56	.26
Creativity level						
Men	77	11.03	3.41	13.07	3.77	3.58**
Women	28	11.96	2.90	14.04	2.98	2.97*
Control	29	13.88	3.76	13.22	3.29	.81
Asymmetrical Preference Test						
Men	77	9.82	4.72			
Women	28	10.00	4.14			.19

* $p < .01$.

** $p < .001$.

the method may be considered to have been even more successful with the Middle East subjects, inasmuch as gains in number of responses were achieved by both sexes in the experimental group; whereas the gain in this variable had been significant only for the women in the American group. Incidentally, subjects in the San Francisco study produced approximately 75 per cent more responses in each phase of the experiment than did the Middle East subjects; a difference that is understandable when one considers that the Americans were operating in their primary language.

Table 2 shows that the significant and positive relationship between APT scores and level of creativity found with the San Francisco subjects does not hold up well in the Middle East. Positive correlations of any appreciable

TABLE 2
CORRELATIONS OF ASYMMETRICAL PREFERENCE TEST (APT) SCORES WITH
NUMBER OF RESPONSES AND LEVEL-OF-CREATIVITY SCORES

Variable	Men	Women
Number of responses		
Phase I	.12	.26
Phase III	.13	.38*
Creativity level		
Phase I	.03	.42**
Phase III	.13	.22

* $p < .05$.

** $p < .02$.

magnitude appear only in the results for women, and two of these fall below the .05 level of significance.

D. DISCUSSION

The positive results obtained as a result of group brainstorming imply that the stimulation of individual creative thinking by such methods may be possible even when participants have varied cultural backgrounds. It may be that some general principles of group learning work to produce this effect: for example, it is possible that the group-brainstorming session gave participants an opportunity to learn how their assigned task might be performed more effectively. It may also be that group brainstorming boosted the interest level of the task. The fact that the level of creativity of the control group declined slightly between Phases I and III may be due to flagging interest. In other words, whatever the control group gained from practice was cancelled by motivational loss.

Although the findings of the present study may be considered encouraging for proponents of group brainstorming, particularly for those who might be considering its use in "developing countries," several limitations should be noted. The fact that group brainstorming facilitated creative thinking in English does not necessarily mean that it would have a similar effect in other languages. Although a relaxed, informal atmosphere is fundamental to successful group brainstorming, it may be difficult to create such an atmosphere—at least among people who are not already on an intimate basis socially—especially if communication takes place in the Mediterranean languages (including Arabic), because those languages place great stress on the formal, syntactical aspects of communication. Some Mediterranean languages do not have any words that convey the sense and feeling of the English word "informal." Whereas English-speaking people (especially

Americans and Canadians) cultivate an informal way of speaking to put others at ease, informality with strangers is not considered to be a desirable thing by Mediterranean peoples, and their languages are not structured in ways that encourage such a relationship. It may be, therefore, that group brainstorming proceeds more effectively in English, even with those for whom English is a second or a third language.

Languages in which the formal aspect rather than the semantic aspect is stressed may create a second kind of problem. A preoccupation with finding the exact or most elegant turn of expression encourages what Guilford (3) terms "convergent thinking," whereas the success of brainstorming depends on the ability to speculate and ramble over a vast range of ideas—i.e., an ability that Guilford calls "divergent thinking." Brainstorming experiments in Mediterranean languages should be undertaken to determine whether or not languages that emphasize formality and structure interfere with the development of divergent forms of creative thinking.

Although the present research presents positive findings regarding the value of group brainstorming to facilitate that type of individual brainstorming concerned with verbal problems (like the writing of cartoon captions), it leaves unanswered the question of whether or not creative solutions of mechanical and technical problems or problems involving broad social implications can be facilitated by group brainstorming.

The present research raises some question about the generality of the relationship between the more-or-less asocial self-concepts of persons preferring asymmetrical art (as measured by the APT) and creative ability. Correlations between the APT and level of creativity in the San Francisco study ranged from .24 to .36, and all were significant at the .05 level of confidence. The failure of similar results to appear in this study may imply that whatever attitudes are measured by the APT interact with creativity only in American culture. This conclusion is confounded by the fact that correlations between various measures of creativity and the APT approach significance for Middle East women.

This latter observation led the authors to determine whether or not some kind of relationship between APT and creativity was being masked by the virtually zero correlations obtained for men. The analysis produced two findings: the first is that men scoring in the middle range of the APT tended to make greater gains in level of creativity than those scoring high or low. The mean gain for those men scoring in the middle third of the APT was 3.77 on the level-of-creativity scale, as contrasted with 1.23 for the bottom third and 1.31 for the top third. The difference between the mean level-of-creativity

score for this middle third and that of the top third and bottom third combined was significant at the .01 level of confidence. This may mean that a moderate degree of the kind of nonconformity the APT measures interacts with group brainstorming to facilitate creative thinking among Middle East men—with a high degree of conformity or nonconformity being less conducive to a positive response to this kind of treatment.

The second finding produced by the additional analysis is that the distribution of male scores on the APT was bimodal, with one mode near the bottom and another near the mean. This type of distribution did not appear in the scores for women. Such bimodality suggests that the APT scores for a number of the male subjects may be understated. Such an interpretation makes sense in view of the fact that there is a general tendency for Middle East men to inhibit the expression of any self-derogatory attitudes, such as the kinds of attitudes (bitter, temperamental, pleasure-seeking, sarcastic, and the like) that are scored "plus" on the APT key. Fake-good tendencies exist with American subjects, of course, but the cultural bans against a man admitting he has such feelings are considerably stronger in the Middle East. Furthermore, being frank, especially being frank to psychologists, is much more of a "virtue" in the United States than it is in the Middle East.

In any event, serious question is raised about the desirability of using the APT as a predictor of creativity for Middle East men. Either the relationship observed in an American sample does not extend beyond the cultural boundary of the United States or a sizable proportion of Middle East male university students cannot be depended upon to give frank and self-revealing responses on a personality questionnaire of the type used.

E. SUMMARY

Middle East university students for whom English is not a primary language participated in a three-phase experiment in which they wrote English captions for cartoons first alone (individual brainstorming), then in a small group (group brainstorming), and finally alone (individual brainstorming). Both men and women displayed significantly more creativity in the third phase, as compared with the first phase, using number of responses and ratings of level of creativity as measures of creativity. A control group did not show any such gains. These findings generally are consistent with those of a similar study conducted with college students in the United States and imply that group brainstorming may function as a facilitator of creativity in cultures other than American, although research should be done to determine whether or not brainstorming would succeed in languages other than English.

Measures of creativity failed to correlate significantly with scores made by men on a self-report, forced-choice questionnaire scored in terms of choices made by persons preferring asymmetrical to symmetrical art; although correlations approached significance for women. Significant and positive correlations between scores on this questionnaire and level of creativity had been found for both sexes in the United States.

REFERENCES

1. BARRON, F. *Creativity and Psychological Health*. Princeton: Van Nostrand, 1963.
2. DUNNETTE, M. D., CAMPBELL, J., & JAASAD, K. The effect of group participation on brainstorming effectiveness for two industrial samples. *J. Appl. Psychol.*, 1963, 47, 30-37.
3. GUILFORD, J. P. *Personality*. New York: McGraw-Hill, 1959.
4. LINDGREN, H. C., & LINDGREN, F. Brainstorming and orneriness as facilitators of creativity. *Psychol. Reports*, 1965, 16, 577-583.
5. OSBORN, A. F. *Applied Imagination*. (Rev. ed.) New York: Scribner, 1957.
6. TAYLOR, D. W., BERRY, P. C., & BLOCK, C. H. Does group participation when using brainstorming facilitate or inhibit creative thinking? *Administrative Science Quarterly*, 1958, 3, 23-47.

QUANTITATIVE ASPECTS OF THE GROUP-PSYCHO-
THERAPIST'S ROLE BEHAVIOR:
A METHODOLOGICAL NOTE*¹

Research Center for Group Dynamics, University of Michigan

JEROME RABOW

A. INTRODUCTION

One crucial requirement for a critical evaluation of any treatment program is the description and measurement of the process of treatment. This description and measurement becomes especially important when group treatment is applied because the focus and manipulation of multiple variables is integral to any dynamically oriented group approach. The complexity of such manipulation suggests the importance of detailed and systematic observation; yet most investigations of group-treatment approaches occur on clinical and descriptive levels (2). Also, these descriptions tend to be generalized rather than specific (6, 7, 8). Any attempts, therefore, to observe in a systematic, parsimonious, and reliable way the extent to which the group therapist conforms to principles of treatment will be valuable for comparing the extent to which group therapists differ and concur in their treatment practices. This note represents such an attempt.

This paper examines some quantitative aspects of role behavior of a group psychotherapist who operates according to the principles of Guided Group Interaction. Elsewhere, the treatment system (4) and principles of this particularized form of group psychotherapy have been stated in detail (1, 3, 9, 10). The focus of the present paper is to suggest that quantitative parameters of therapist's behavior may indicate the extent to which practice proceeds according to particular principles of treatment.² It may signify the importance of the quantitative aspects of the group psychotherapist's role behavior as a variable for detailed study.

* Received in the Editorial Office, Provincetown, Massachusetts, on January 17, 1964. Delayed one issue for technical reasons. Copyright, 1965, by The Journal Press.

¹ This study was conducted as part of an overall research design to evaluate a community-based delinquency-rehabilitation program in Provo, Utah. The author is indebted to LaMar T. Empey and Maynard Erickson of the Provo Experiment for their support and aid in the formulation of the research and to Martin Gold of the Research Center for Group Dynamics, University of Michigan, for his constructive suggestions. Support by the Ford Foundation made this evaluation possible.

² A later study will report on qualitative aspects of role performance.

B. QUANTITATIVE ASPECTS OF ROLE PERFORMANCE

One of the specific principles of Guided Group Interaction revolves around the activity of the group leader. Since the theoretical assumption for causation involves the view that most delinquent behavior is purposive and adaptive, the therapist (rather than waiting for the offender who is committed to his own way of life to evaluate his own motives and values) should actively involve himself in posing alternatives and in challenging the neutralizations and rationalizations used by offenders (5, 11). These efforts at posing alternatives have to be interwoven with efforts to create an atmosphere in which boys feel free to discuss their behavior. Since the therapist also attempts to steer the group into discussions that he feels have greater therapeutic potential, his activity tends to decrease as the group develops and begins to discuss and focus on the areas to which the therapist has guided its members. A first hypothesis, then, is that the therapist operates at a more active rate in the earlier meetings than in the later sessions; also, that the total number of statements he makes in the earlier meetings is greater than the number of statements he makes later on. Two aspects, then, of therapist behavior studied are the rates of interaction and the total number of statements per session.

One other aspect of the quantitative role that the group leader plays in Guided Group Interaction concerns the type or nature of his statements. In the initial stage of group development, the leader's statements are characterized by their "challenging" nature. These challenges or interventions are derived from the view that the type of delinquent boy involved has a purposive and adaptive outlook. The view is that the delinquent is a boy who has rarely stopped to consider his own motives. Even if he has, like other delinquents he is equipped with a series of rationalizations that successfully justify delinquency, serve to disparage his rejection by representatives of conventional society, and counter whatever conventional norms the delinquent may have internalized (11). Examples of these challenges are: "Why did you do that?" or "What do you mean when you say you thought you wouldn't be caught?"

Other statements in the early stages are supportive of the group and of individual efforts by the members. A typical supporting statement might be: "Joe really told us how he felt tonight," or "Why was the group able, today, to make progress with Pete's problem?" Both the supporting and challenging statements tend to be brief and compact in nature.

In a later stage of group development, the therapist tries to make members aware of discrepancies between behavior and stereotypes and aids in articulating unconscious motivation. Also, the group leader can provide terminology

and frames of reference for classifying individual behavior. Since the therapist's role changes from an active directive, challenging, and supportive role to a less active and less directive interpretative role, the statements he makes in the early sessions tend to be shorter and more compact than statements made in the later stages of group development. This aspect of the therapist's behavior was also evaluated and is the third focus of this study.

C. METHODOLOGY

The study described here is based upon a 10 per cent sample of all recorded statements made by the therapist. The therapist whose behavior was examined worked with a group of boys who met for a total of 76 sessions. These sessions were held over a period of four months, and the number of meetings used in the analysis numbered 66. Six meetings were not recorded due to mechanical difficulties with the recording equipment, and four meetings were not used because the group leader, whose behavior is the focus of the study, was not in attendance.

Through the use of a table of random numbers a random start was made in each of the 66 meetings, and every 10th statement that the therapist made was transcribed and coded.

A statement was defined operationally as a remark made by the therapist that communicated some idea or thought or asked a question.

Remarks that communicated nothing beyond the fact that the therapist heard the speaker's comments were not included in the sample. Remarks by the therapist indicating that he had heard the speaker and wished him to continue or to repeat his comments were also excluded. Examples of these remarks are: "Yeah," "Uh-hum," or "Will you repeat that?" If a statement was inaudible because of recording difficulties or interrupted by another group member, it was also omitted from the sample. Of the 41 statements that were not utilized because of the foregoing criteria, 29 were interrupted statements, nine were inaudible, and two were statements that were repeated. In each of the 41 cases, the next remark by the therapist was selected for analysis. All of the uncounted, interrupted, inaudible, and repeated statements occurred within the first 33 meetings. This fact suggests that after a certain stage of development group members tended to allow the therapist to complete his remarks rather than interceding in attempts to deflect intervention. Communication between therapist and group members seemed to improve with the group's development, as members tended to interrupt less frequently. The 10 per cent sample of statements for the 76 meetings comprised 1074 statements.

Sampling accuracy was diminished by three factors. (a) All of the therapist's statements were not recorded. After one hour a tape had to be turned over; so about 30 seconds of meeting time went unrecorded. (b) Ten meetings were not used in the analysis. (c) Occasionally, the boys would start the meetings without waiting for the observer who was responsible for starting the tape. These types of errors are small, they occurred on a random basis, and were relatively infrequent.

D. RESULTS

In this study, with its emphasis upon quantitative parameters, the author did not test hypotheses in any strict sense. Statistical measures of description and inference are reported in the hope that they will promote interest in comparative measurements for group therapists of differing orientations and generate hypotheses that might be tested in some future work.

1. Number of Statements

The data indicate that the group leader made an average of 162 statements per meeting. Because group sessions lasted approximately 75 minutes, the therapist participated at the approximate rate of two statements per minute [see Figure 1]. During the first 33 sessions, the mean numbers of statements made by the therapist was 20.7; during the second set of 33 sessions, it was 11.8. The difference between these means is significant at the .01 level. A

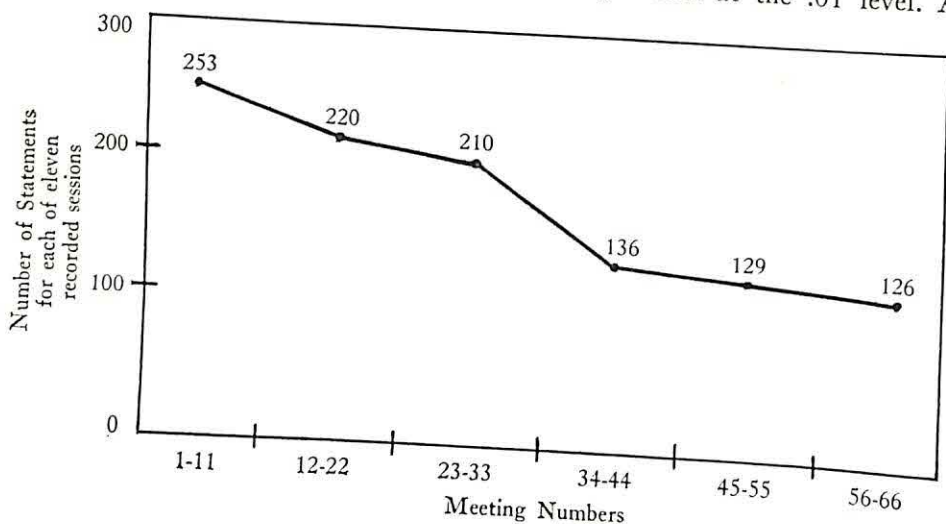


FIGURE 1
NUMBER OF STATEMENTS MADE BY THERAPIST

Spearman rank-difference correlation (using meeting number with correction for ties) is $-.74$. Kendall's tau is $-.53$. Thus, the longer the therapist meets with his group the less active he becomes. This result indicates conformity to the role prescriptions of Guided Group Interaction.

2. Rates of Interaction

To determine the rate at which the therapist interacts with group members, the amount of footage between every 10th statement was computed.

The average amount of footage between randomly selected statements for the 66 meetings was 54.1. In interaction rates, the therapist made a 10th statement every 268 seconds or every half minute.

When the average rates of interaction in each of the 66 meetings are compared with the average rate of interaction for all 66 meetings, one discovers 17 meetings for which the rates of interaction fall within two standard deviations of the average interaction rate for all 66 meetings. Table 1 shows

TABLE 1
THE NUMBER OF SIGNIFICANTLY HIGH³ AND
SIGNIFICANTLY LOW INTERACTION RATES

Interaction rate	Meetings		Total
	1-33	34-66	
Significantly high	19	2	21
Significantly low	2	26	28
Total	21	28	49

that therapist interaction is significantly higher in the first half of the group sessions as compared with that in second half of the group sessions. The chi square of 33.9 is significant beyond the .001 level. The data demonstrate that the therapist takes a more active role initially in the group meetings and decreases his participation as the group develops.

3. The Number of Words in a Statement

The average number of words for each statement is 10.5. This indicates that the therapist tends to make statements that are short and compact. These statements are not interpretations or explanations, but statements that question or speculate about an individual in the group. The therapist's statements might be seen as stimuli that other group members might pick up and develop and utilize for purposes of making their own interpretations or to test their own reality perceptions of the situation.

³ Significant is used in the statistical sense of the number of words falling above and below two sigmas from the mean.

The reader may recall that the role prescription for the group leader in the initial stages of Guided Group Interaction prescribes that the therapist challenge the assertions and rationalizations of the group members. Thus, the group leader, if he is fulfilling the role prescriptions, tends to utilize other group members for the purpose of making interpretations. As the group develops into an effective treatment unit, the therapist can make interpretations like other group members. This latter type of statement tends to be longer than remarks designed to challenge group members. As suggested earlier, it might be expected that the length of statements increases as the meeting develops.

Table 2 indicates the meetings in which the therapist made statements that are significantly longer and shorter than the overall average of 10.5 words per statement.

TABLE 2
THE NUMBER OF SIGNIFICANTLY LONG AND
SIGNIFICANTLY SHORT STATEMENTS

Statements	Meetings		Total
	1-33	34-66	
Significantly long	8	20	28
Significantly short	19	5	24
Total	27	25	52

The chi square for Table 2 is 12.1, significant beyond the .001 level; thus, there is a significant association between the length of the therapist's statements and the length of time that the group has met.

E. SUMMARY AND CONCLUSION

In assessing the role prescriptions of a group-therapist activity, the number of statements per meeting, the rates of interaction per session, and the length of statements may be considered important parameters of behavior. A straightforward counting process was used to study and evaluate the foregoing dimensions of the behavior of one therapist. Role descriptions for the therapist in Guided Group Interaction were described and evidence for adherence to role requirements was presented.

The attempt to measure the role performance of a group therapist who is guided by particular treatment principles shows that all the quantitative role prescriptions were followed. Both the number of statements and the rates of interaction decreased significantly over time. Furthermore, the length of statements increased significantly during the later therapy sessions.

In the group observed, the therapist adhered to the limited quantitative

treatment principles specified by the theory of treatment. This note suggests that the attention to quantitative parameters may be one method by which systematic differences between therapeutic schools may be discovered.

REFERENCES

1. ABRAHAM, J., & McCORKLE, L. W. Group psychotherapy on military offenders. *Amer. J. Sociol.*, 1946, **51**, 465-469.
2. BENNIS, W. G. A critique of group theory research. *Int. J. Group Psychother.*, 1960, **1**, 63-77.
3. BIXBY, F. L., & McCORKLE, L. W. Guided group interaction in correctional work. *Amer. Sociol. Rev.*, 1951, **16**, 455-459.
4. EMPEY, L. T., & RABOW, J. The Provo experiment in delinquency rehabilitation. *Amer. Sociol. Rev.*, 1961, **26**, 679-695.
5. GLADSTONE, H. P. A study of techniques in psychology with youthful offenders. *Psychiatry*, 1962, **25**, 147-159.
6. GLATZER, H. T., & DURKIN, H. E. The role of the therapist in group relations therapy. *Nervous Child*, 1945, **4**, 243-251.
7. HAGAN, M., & KENWORTHY, M. The use of psychodrama as a training device for professional group workers in the field of human relations. *Group Psychiatry of Socio-Psychopathology and Society*, 1951, **1**, 23-40.
8. KONOPKA, G. The role of the social group worker in the social setting. *Amer. J. Orthopsychiat.*, 1952, **22**, 138-142.
9. McCORKLE, L. W., ELIAS, A., & BIXBY, F. L. *The Highfields Story*. New York: Holt, 1958.
10. RABOW, J. The role of the therapist in guided group interaction: A description of an applied treatment approach to habitual offenders. Unpublished Master's thesis, Columbia University, New York, 1962.
11. SYKES, G., & MATZA, D. Techniques of neutralization: A theory of delinquency. *Amer. Sociol. Rev.*, 1957, **22**, 664-670.

University of California, Los Angeles
Sociology Department
Los Angeles, California

A DELAY-OF-REINFORCEMENT EFFECT IN PERSUASIVE COMMUNICATION?*

Communication Research Division, Boston University

RALPH L. ROSNOW¹

A. PROBLEM

In a recent study by Rosnow and Russell (3), opposing persuasive communications were read to high-school students either just prior to the onset of a "reinforcing" event or immediately following that event. When communications preceded reinforcement, findings supported an hypothesis derived from the earlier Law of Effect; i.e., respondents' opinions were found to have been significantly changed in the direction advocated by the one communication (either pro or con) that was closer in time to the reinforcement. A similar association, based on a spread-of-effect hypothesis, had been posited for a case when communications followed reinforcement. These results were not statistically significant ($p > .05$), and this second hypothesis was rejected.

In that study as well as in one by Rosnow and Lana (2), "reward" was characterized as the relief that ensued after release from a punishing situation. Findings lent support to the contention that a tendency to respond in the directions advocated by the communications would be strengthened for the one communication preceding and closer in time to the rewarding event. When communications followed that rewarding event, it has been suggested (3) that since communications are not instantaneous stimuli, a temporal lag must exist between the onset of the stimulus complex (the communications) and the opinion reaction. When communications preceded reward, this reaction led into the rewarding event. However, when communications followed a rewarding event, the delay in opinion reaction time may have far exceeded what would have been optimally appropriate to link the communication nearer to the rewarding event with that event, due to "decay" in the reward gradient. The connection would have been expected, therefore, to be weak; and a nonsignificant order effect would appear plausible.

* Received in the Editorial Office, Provincetown, Massachusetts, on March 25, 1964. Copyright, 1965, by The Journal Press.

¹ I am grateful to Sarah Tanenbaum and Richard Gaskill of Wakefield High School, Virginia, and to Gordon W. Russell of Juniata College for their generosity of time and for their skilled assistance in running subjects; and to Robert E. Lana and Edward J. Robinson for valuable criticisms of an earlier version of this paper.

The present paper is a report of an experiment designed to explore the effect of a "punishing" event on opinion change, when that event was either closely preceded in time or closely followed in time by opposing persuasive communications. Following the rationale of the two other studies (2, 3), it was now contended that a tendency to respond in the directions advocated by the communications would be weakened for the one communication closer in time to a punishing event. When both communications preceded punishment, it was hypothesized that any opinion change would be in the direction of primacy (i.e., the direction advocated by the first of the two opposing communications, the one farther in time from the punishing event). When both communications followed punishment, this rationale would now suggest, if anything, a weak recency effect (i.e., opinion changed in the direction advocated by the last communication, the one farther in time from the punishing event).

B. PROCEDURE

Ss were 84 third-year and fourth-year high-school students in four psychology classes, all of whom had no prior exposure to this type of experiment.

Pretest opinion questionnaires were administered by *E* during a regularly scheduled class meeting, at which time he was introduced as someone from the University who is "polling students' opinions on the United States' use of nuclear weapons."

Questionnaires were completed, following which *E* thanked the instructor and the students for their cooperation and left the classroom. One week passed between pretest and treatment, thus reducing the effect of recall of questionnaire responses and the possibility of pretest sensitization as complicating variables. On the seventh day, the instructor began two of the four classes (A_1 and A_2) by expressing his disappointment over the quality of the students' homework and announced a surprise quiz to "see how much you really know." The quiz consisted of two open-end questions related to topics covered in class but the content of which was of an esoteric nature, and the students were permitted only five minutes to record their answers. Immediately upon having collected their papers, the instructor reintroduced *E* who then read pro and con communications² to Ss (using reversed orders of presentation for the two classes).

² Copies of the persuasive communications, which were taken from Lana's research report (1), may be obtained from the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Order Document No. 7481, remitting \$1.25 for photocopies or 35-mm. microfilm.

In the remaining two classes (A_3 and A_4), E read the communications immediately prior to announcement of the "punishing"³ quiz. Posttest opinion questionnaires were completed by all S s after completion of the punishing event and reading of the communications.

A fifth group of students, using a Likert-type questionnaire, rated the communications for direction and strength and rated the topic for its perceived controversialism.

C. RESULTS

The questionnaires and communications were the same as those used in Rosnow and Russell's study. The pro and con communications were around 425 words long and took a little over four minutes to present. Pretest and posttest opinion questionnaires were identical, each consisting of five five-point Likert-type items. The range of scores was five through 25, with five indicating the most unfavorable attitude and 25 indicating the most favorable attitude toward "use of nuclear weapons."

For the group of students who rated the communications for direction and strength and the topic for its perceived controversialism, results were consistent with previously reported findings (3). "Use of nuclear weapons" was rated as moderately controversial in nature, while the pro and con communications were rated as opposite in direction; but were not rated as significantly different from one another in terms of the perceived strength of their content ($df = 21$, $t = .15$, $p > .05$).

The subtractive procedure that results in t tests for differences between pretest and posttest opinion scores for the four treatment groups confirms the primacy prediction but not the recency prediction. Primacy had been predicted when communications preceded a punishing event, and results supported this expectation ($df = 33$, $t = 2.1$, $p < .05$). When communications followed punishment, the results suggested recency but fell at a nonsignificant level ($df = 47$, $t = .3$, $p > .05$).

D. DISCUSSION

Whereas for Thorndike the spread of effect represented a gradient of reward, there has been evidence by Tilton (4) to suggest an analogous spread of negative effects or gradient of punishment. It was from this latter position

³ Evidence supporting interpretation of this event as "punishing" may be found in the Rosnow and Lana study (2) where mood questionnaires were completed by subjects just before and two days after a difficult surprise quiz. Findings suggested that anxiety was effectively induced and maintained as a result of taking the quiz.

that the second of our two hypotheses, the one predicting recency, was derived.

The predicted direction was, in fact, obtained; but the effect was too weak to be taken as evidence for a "spread." This still does not rule out the possibility of obtaining such a spread of effect in this kind of persuasive-communication situation, but it does suggest the necessity for either some modification in the length of the communications or for more sensitive measures of change in the immediate region of the reinforcement gradient.

On the other hand, hypotheses derived from the earlier Law of Effect have been consistently confirmed (both in the case of reward as well as punishment). Previously, Rosnow and Russell reported that respondents' opinions were changed in the direction advocated by the one of a pair of opposing arguments more immediately preceding a rewarding event. This suggested that contiguity between the second-presented argument and the reward strengthened the tendency to respond in the direction advocated by that argument. Now we find that opinions are changed in the direction advocated by the one of a pair of opposing arguments preceding (but farther in time) from a punishing event. This suggests that contiguity between the second-presented argument and the punishment has weakened (or perhaps suppressed) any tendency to respond in the direction advocated by that second argument; this has permitted the first argument to have the greater influence on opinion change.

The unique nature of what have been characterized as rewarding or punishing events has produced a situation where there is a temporal lag between presentation of the persuasive messages and either onset (when reinforcement preceded the communications) or completion (when the communications preceded the reinforcement) of that event. Both the present study as well as the one by Rosnow and Russell suggest a typical delay-of-reinforcement gradient.

E. SUMMARY

After reading opposing persuasive communications to Ss just prior to the onset of a "punishing" event, opinion was found to be significantly changed ($p < .05$) in the direction advocated by the one communication (either pro or con) farther in time from that event. When communications followed punishment, results (though not statistically significant) suggested a similar trend. Findings supported an hypothesis derived from the earlier Law of Effect but did not confirm a spread-of-effect hypothesis. The present findings,

considered in the light of a study reported earlier, suggest a typical delay-of-reinforcement gradient.

REFERENCES

1. LANA, R. E. Order effects in persuasive communications. Progress report No. 1 to the National Institute of Mental Health, USPHS Grant M-4830, May, 1962.
2. ROSNOW, R. L., & LANA, R. E. Complementary and competing-order effects in opinion change. *J. Soc. Psychol.*, 1965, **66**, 201-207.
3. ROSNOW, R. L., & RUSSELL, G. Spread of effect of reinforcement in persuasive communication. *Psychol. Reports*, 1963, **12**, 731-735.
4. TILTON, J. W. The effect of "right" and "wrong" upon the learning of nonsense syllables in multiple choice arrangement. *J. Educ. Psychol.*, 1939, **30**, 95-115.

Communication Research Division
Boston University
640 Commonwealth Avenue
Boston, Massachusetts 02215

THE EFFECT OF "HUMAN" VS. "MACHINE" SET ON GROUP PROBLEM-SOLVING PROCEDURES*¹

Educational Testing Service, Princeton, New Jersey

JOHN K. HEMPHILL AND CAROLYN B. McCONVILLE

A. INTRODUCTION

Much research has been conducted that has attempted to contrast the effectiveness of the individual working alone with the effectiveness of groups in problem solving (1). The results of much if not all of this and similar work may have more meaning as more is understood about the basic characteristics of group tasks and how these characteristics are related to problem-solving processes. Concepts need to be introduced into research on group problem solving that are oriented toward the description of structural task characteristics.

Progress has been made recently toward an analysis of some of the demands tasks make upon the interacting relationships among group members. One important step was the recognition of the distinctions between disjunctive as contrasted with conjunctive task requirements (10). If the relationship between the acts of group members and their progress toward goal achievement is disjunctive (because of the "parallel" relationship of subtasks), the individual's contribution to task achievement can be made independently of the acts of other members of the group. On the other hand, if relationships between the acts of individual group members and achievement of the group's goal are conjunctive, then what the individual does may or may not be effective, depending completely upon what some other member of the group has done or will do in the future. Clearly, group tasks having disjunctive requirements and those primarily of the conjunctive variety may call for radically different procedures for their solution.

The experiment to be described is a study of the effects of two contrasting "sets" introduced to the subject about the nature of the "others" with whom he was working, where the total task involved predominantly conjunctive connections between the subtasks. In the case of the "human" set, the subject

* Received in the Editorial Office, Provincetown, Massachusetts, on March 26, 1964. Copyright, 1965, by The Journal Press.

¹ This research was supported by a contract between the Office of Naval Research (Group Psychology Branch) and Educational Testing Service.

was led to believe that he was working with another human subject on the task. In the case of the "machine" set, the "other" was represented as a machine that had been programmed to respond in a complex manner to the acts of the subject. In the case of both sets the response actually was made by the "experimenter," who followed a predetermined structure that was identical for the two conditions.

The concern of the investigation is that subjects under the influence of the two sets may make different responses to the objectively identical structure they find in response to their acts.

From a point of view suggested in a theory of leadership and group behavior (3, 4, 8), the individual responds to working with others on mutual problems by noting consistencies in their responses and by adjusting his own behavior in a manner that meets the required interlocking of behavior. It was also suggested that the individual might tend to introduce consistency in his own behavior as a means of making the "other's" task of predicting his behavior more tractable. Thus, if the individual is working on a mutual task with one or more persons (and assuming high motivation), his success would depend on (a) the accuracy of his analysis of the task requirements, (b) his understanding of the probable behavior of the others, (c) his formulation of an adequate course of action to meet the task requirements, and (d) his communication to the others of his intended behavior. If the individual is set to regard the task he faces as "mechanical" and as not involving the behavior of some other individual, his success depends only on the accuracy of his analysis of the task requirements and on his formulation of an adequate course of action to meet the task requirements. With this latter set he has no concern with the probable behavior of another person or with the problem of communicating his intentions to someone else. The differences in problem-solving strategy that the individual may develop under these "two" contrasting sets may be small, however. The technique or procedure by which the individual can secure information about task requirements and by which he can learn about consistency in the behavior of other persons with whom he is interacting may be very similar.

The general hypothesis of the experiment may be simply stated as follows: Individuals attempting to solve a task involving a conjunctively connected subtask will approach their work on the task in a different manner depending upon whether they are set to interact with a mechanically responding environment or are set to interact with an environment that includes human responses. In other words, the problem solver, when working with other humans, will modify his behavior to take into account human characteristics of his partner.

Rejection of this hypothesis cannot be interpreted as suggesting that social factors in group problem solving are unimportant, but may lead to specifying and examining conditions under which differences in the individual's approach to a group task might be better understood.

B. PROCEDURES

Twenty-four men and 24 women, matched only on educational level, served as experimental subjects. The men represented all undergraduate classes at Princeton University, and most of the women were wives of graduate students at the same institution. All had had some college education, but none had advanced degrees. No attempt was made to match subjects within pairs.

The 48 subjects were run in pairs, women subjects separately from the men. Instructions were given to 12 pairs of the subjects (six of each sex) at the beginning of the experiment in a manner calculated to lead them to believe that the two members of the pair would work together as partners on the tasks they were to do. The remaining 12 pairs were led to believe that they would work as individuals and would respond to a complex machine. Each subject worked on two distinct tasks (I and II) that were counter-balanced throughout the experiment in ABBA order.

The two subjects were seated in a room partitioned by a folding door, giving privacy to each individual. One-way vision mirrors permitted the experimenters to observe the subjects as they worked. On the table in front of each subject was an electrical plugboard, arranged in six rows and five columns of one and one-quarter inch squares; each square, letter-number coded (E1) for easy identification, contained a telephone jack and a neon light. Figure 1 is a simple sketch of one of the boards.

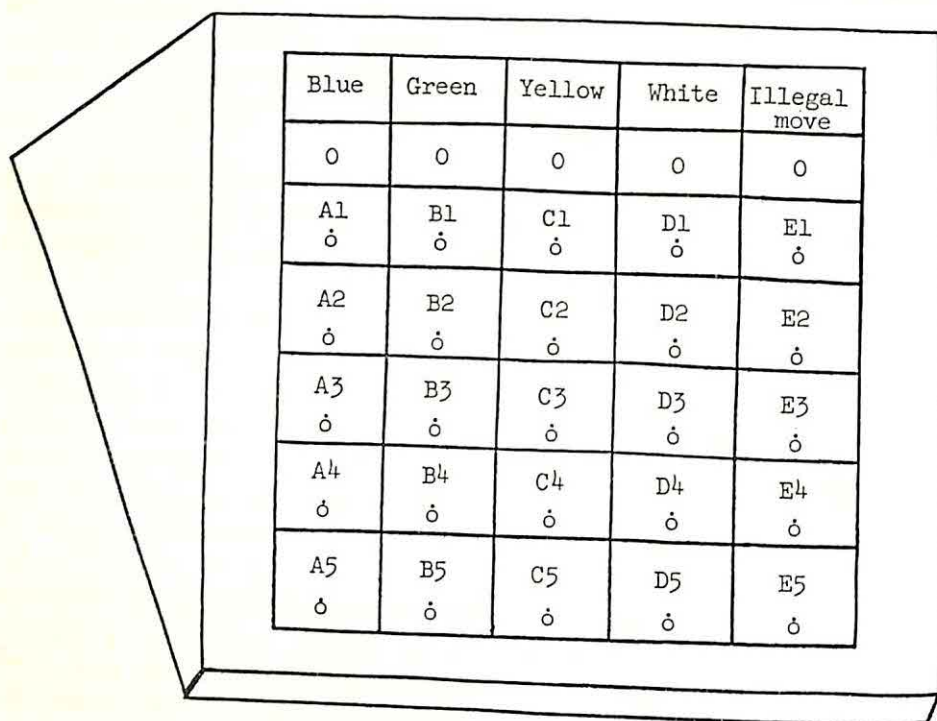
Each subject was supplied with a set of four telephone plugs to be used with his board. Each plug was marked with a different color: blue, green, yellow, or white. Two of the plugs were wired internally in a manner such as to cause the neon light in a square to come on when it was inserted into the jack in that square. The other two plugs were dummy plugs and did not affect the lighting of the squares.

Instructions given to the subjects at the beginning of the experiment were used as the means of establishing a "human" or a "machine" set toward their partners. In order to make the sets more plausible, it was necessary to include the following elements in the experimental situation. (These elements, although trivial in themselves, led to seemingly important differences in experience with the tasks.)

To create the "human" set, a subject was told that his partner was working

at an identical board and that his partner also had two "active" plugs and two "markers," but that his active and marker plugs were of the opposite colors—thus *Subject A* controlled the blue and yellow active plugs; *Subject B*, the green and white ones. Each subject was permitted to move one plug at a time, taking alternate moves.

The "machine" set was created by instructions explaining that the boards were connected with an electrical switching mechanism and that the "machine"



Blue	Green	Yellow	White	Illegal move
O ○	O ○	O ○	O ○	O ○
A1 ○	B1 ○	C1 ○	D1 ○	E1 ○
A2 ○	B2 ○	C2 ○	D2 ○	E2 ○
A3 ○	B3 ○	C3 ○	D3 ○	E3 ○
A4 ○	B4 ○	C4 ○	D4 ○	E4 ○
A5 ○	B5 ○	C5 ○	D5 ○	E5 ○

FIGURE 1
A SKETCH OF THE ELECTRICAL PLUGBOARD USED IN THE EXPERIMENT

would respond to the manner in which the subject moved his active plugs by selecting and lighting a square on the following move; thus the subjects and the "machine" worked together, again taking alternate moves. *Both Subjects A and B controlled the blue and yellow active plugs and made their moves at the same time.* Since the subjects came to the laboratory in pairs, it was stressed that the apparatus could accommodate two people; but that they would be working entirely independently.

The goal of each of two tasks that were given to the subjects was the same; i.e., starting from positions in the lower left-hand corner of the board the

plugs were to be moved according to specific rules in a manner that one of them would light the square in the uppermost right-hand corner of the board. The object was to achieve the goal in as few moves as possible. The plugs were permitted to be moved only in a specified color order, and the four plugs had to remain together as a group, maintained at all times by a horizontal or vertical connection.

At no time did the subjects interact with each other, as they believed they were doing in the "human" condition. In reality, each subject's board was interconnected to a separate board controlled by one of two experimenters; thus there were two independent sets of two boards in series. The experimenters followed a predetermined structure that was identical for the two conditions.

In the instructions to the subject, he was asked to predict either his "partner's" next move or the "machine's" next move depending on the experimental condition in which the subject was participating. Comparable mimeographed forms were provided for this purpose.

Each subject was given eight trials per task or 16 trials in the total experiment. Each trial consisted of 10 moves. If the goal was reached in less than 10 moves, both experimenters and subjects replaced their plugs in the occupied squares for the duration of the trial. For a perfect solution, Task I required eight moves; Task II, seven moves. The experimenter's play was governed by rules that defined a pattern of moves that was effective in reaching the goal square in the minimum number of moves, *providing* the subject made the appropriate move at each opportunity.

A postexperimental questionnaire was administered to determine attitudinal differences on the part of the subjects participating in the two conditions. The instrument was patterned after Osgood's method of semantic differentials and ascertained attitudes pertaining to (a) the experiment as a whole, (b) the subject's own play during the experiment, and (c) the responses made to his play (7). Most of the word pairs reflected common stereotypes of differences between machine and human performance.

C. RESULTS

The data provided by the subjects in completing this postexperimental questionnaire were examined by means of two 10-variate multivariate analyses of variance (2).² One of the two analyses concerned 10 semantic-differences

² The authors express appreciation to Professor R. Darrell Bock of the Psychometric Laboratory of the University of North Carolina for suggesting the multivariate analysis of variance and for supervising the analysis of the data.

tial scales that the subject had used to describe his reaction to the experiment in general. The second of the two analyses was of five scales that referred to the subject's views of his acts and five scales that referred to his views of the responses he had received to his acts. Two experimental classifications, condition (human *vs.* machine) and sex (men *vs.* women), were considered in the analyses. In the first of the two analyses, the following 10 semantic-differential pairs were included:

- | | |
|---------------------------|------------------------|
| 1. Friendly-Unfriendly | 6. Public-Private |
| 2. Human-Mechanical | 7. Sociable-Unsociable |
| 3. Meaningful-Meaningless | 8. Rhythmic-Irregular |
| 4. Honest-Dishonest | 9. Rigid-Flexible |
| 5. Emotional-Unemotional | 10. Orderly-Confused |

The results of the analysis were not encouraging insofar as the significance of the over-all effect of the "human" *vs.* "machine" conditions was concerned. This effect is not significantly larger than the estimate of experimental error. Sex differences may be present in the manner in which the subject reacted to the experiment (or perhaps in response set to the rating tasks). The significance level for the test (.06) is not small enough to reject the hypothesis of no association with sex differences with great confidence, but produces considerable doubt. The interaction of sex differences and the experimental conditions does not appear to produce a significant effect.

Despite the nonsignificant and doubtful outcome of the over-all effects of the conditions and sex differences, an examination of the 10 variates individually was undertaken with the hope of clarifying some of the ambiguity. Caution must be exercised in interpreting the results, since it has already been shown that the over-all analysis of the variates does not reach the .05 level of significance. However, the *trends* in the detailed data may be worth noting since they follow what might have been expected. Under the "human" condition, the experiment appeared to the subjects to be better described as (*a*) human rather than mechanical, (*b*) more emotional than unemotional, and (*c*) more sociable than unsociable. These seem reasonable contrasts. What sex differences there are appear to involve the tendency of the men subjects to regard the experiment as more private than do the women subjects.

In the over-all analysis of the second set of 10 variates, both "human" *vs.* "machine" conditions and sex differences appear to have significant effects at the .03 and .02 levels respectively. The interaction of the two is not significant.

The variates considered in the second analysis are listed in Table 1. From this table it becomes apparent that under the "human" condition those subjects who had been instructed that they were working with a partner regarded

the responses they received to their acts as less "clear" than did the subjects who had been instructed that the responses to their acts were generated by a "machine." This Clear-Confused Scale was the only variate among the 10 that was clearly related to the "human" *vs.* "machine" conditions. Leading-Following, as applied both to the subject's own acts and the responses to his acts, tends to be the next most sensitive pair. There appears to be a slight tendency for "leading" to be associated with the "human" condition, as applied both to the subject's own acts and to the responses. In other words, the subjects

TABLE 1
DIFFERENCES BETWEEN THE EXPERIMENTAL CLASSIFICATIONS
FOR THE 10 VARIATES OF SET 2

Variate	Human <i>vs.</i> machine			Sex differences		
	$\bar{X}_h$	$\bar{X}_m$	<i>F</i>	$\bar{X}_m$	$\bar{X}_f$	<i>F</i>
Subjects' acts						
Responsive-Unchanging	2.84	3.04	.17	3.21	2.67	1.18
Leading-Following	3.00	3.80	2.99	3.02	3.71	1.86
Emotional-Analytical	5.34	5.46	.11	6.12	4.66	15.44*
Considerate-Inconsiderate	2.88	2.70	.16	2.66	2.92	.37
Cooperative-Competitive	2.88	3.46	1.00	2.79	3.54	1.65
Responses to subjects' acts						
Leading-Following	3.29	4.04	2.63	3.92	3.42	1.17
Cooperative-Competitive	3.04	2.71	.47	2.54	3.21	1.88
Considerate-Inconsiderate	2.92	3.16	.35	2.75	3.34	1.93
Responsive-Unchanging	3.50	3.79	.21	4.12	3.16	2.30
Clear-Confused	3.92	2.46	10.09*	3.21	3.17	.01

* Significant beyond the .01 level.

who worked under the "human" condition viewed their own work as involving "leading"; but also apparently felt that their partners were "leading" too. Perhaps the subjects who worked under the "human" condition were set in a manner that permitted them to rationalize their difficulties with the task by projecting some of their confusion as a characteristic of their partner.

Sex differences in response to the experiment appear to be confined to one variate, Emotional-Analytical. Men subjects, as compared with women subjects, viewed their acts as more analytical or were less willing to describe their work as emotional.

D. QUALITY OF PERFORMANCE

The number of "moves remaining" to be made after the eighth move on Task I or after the seventh move on Task II was used for the purpose of determining the effectiveness of task performance. It will be recalled that eight and seven are the minimum number of moves required to complete

these tasks. "Moves remaining" is a measure of distance between the actual performance and the best possible performance taken after the minimum number of moves had been made that could have completed the task. This distance is expressed in terms of the number of moves that would have to be made to bring the actual performance to the level of the best possible performance. In determining "moves remaining," the position of the square just occupied at the completion of the eighth (or seventh) move was noted, and the number of squares between that square and the goal square was counted with proper observation of the restriction that legal moves can be made only to those squares that border a square on its four sides (city-block distance). This measure reflects inadequate task performance in more detail than does a simple count of perfect trials, in that it differentiates degrees of inadequate performance.

In the analyses of the number of "moves remaining," four separate partitions of χ^2 were made per task (5, 6, 11). First, the subjects were divided into Groups A and B, with 24 subjects each. As pointed out previously, the sets of tasks facing Subjects A and B were not identical nor were they completely comparable between conditions. Lack of comparability appeared with the first move made and, because of the conjunctive nature of the connections between moves, continued throughout the remaining moves. Within each of the groups of 24 subjects, performances on the four trials in which the experimenter made the first move were separated from those in which the subject made the initial move. This separation was necessary because of the differences in objective indices of difficulty associated with whether the subject or the experimenter was first to move on the trial.

It was found that the subjects were most effective in terms of this measure on those trials of Task II where the experimenter moved first and apparently were least effective on Task I where they moved first. More detailed examination showed that Task I had an average moves-remaining score of 1.78 per trial, as compared with 1.11 for Task II; Subjects A averaged 1.33 moves remaining after each trial, as compared with 1.54 for Subjects B; and the average number of moves remaining per trial where the subject moved first was 1.99, as contrasted with .91 where the experimenter was first to move.

Table 2 shows in detail the partitions of χ^2 that were made of the number of moves remaining for Task I, experimenter first, and Subject Group A. With one and 16 degrees of freedom, an F of 4.49 is significant at the .05 level; 3.06, at the .10 level. Only one of the obtained F ratios reached the .10 level of significance, and this F was for the interaction source involving conditions and task order. There is, therefore, a slight suggestion that whether

TABLE 2
PARTITION OF CHI SQUARE ASSOCIATED WITH NUMBER OF MOVES REMAINING
(Task I, Experimenter First, Subject Group A)

Source	<i>df</i>	χ^2	<i>MS</i>	<i>F</i>
1. Between conditions	1	3.409	3.41	2.10
2. Between sexes	1	2.817	2.82	1.74
3. Between task orders	1	4.761	4.76	2.94
4. Interaction (1 x 2)	1	2.817	2.82	1.74
5. Interaction (1 x 3)	1	6.338	6.34	3.91
6. Interaction (2 x 3)	1	2.817	2.82	1.74
7. Interaction (1 x 2 x 3)	1	2.817	2.82	1.74
8. Between individuals	16	25.915	1.62	
Total	23	51.690		

Task II preceded or followed Task I may have been a factor in making a difference between the "human" and "machine" condition. The mean number of moves remaining for the four comparisons involved in this interaction are as follows:

Machine Condition	Task II	Follows Task I	1.83
Machine Condition	Task II	Precedes Task I	.75
Human Condition	Task II	Follows Task I	1.75
Human Condition	Task II	Precedes Task I	1.83

These data suggest the hypothesis that individuals working under the "machine" condition may learn from experience with a previous task, but that under the "human" condition no such learning occurs. This observation appears to be consistent with the attitudinal data in which it was shown that the major difference introduced by the "human" *vs.* "machine" sets was that subjects who worked under the "human" condition tended to regard their partners as *confused*. They might not have expected to learn from the play of the others, and thus may not have been as attentive to this play as were the subjects under the "machine" set.

This detailed discussion of data from the first classification category of Table 3 illustrates the manner in which the *F* ratios may be interpreted. Table 3 summarizes the *F* ratios for each of the eight analyses. A definite tendency appears for some of the tasks to be differentially difficult for the men and women subjects [see Column 2]. In every case the difference was in the direction indicating that the men subjects were able to perform the tasks more effectively. Task order [see Column 3] appears significant at the .05 level for two of the categories of classification of tasks. These two involve Task I when the subject moves first. This difference is perhaps best regarded as a

TABLE 3
F RATIO SUMMARIZING THE RESULTS OF PARTITIONS OF CHI SQUARE OF THE NUMBER OF MOVES REMAINING
FOR EIGHT CATEGORIES OF TASKS, FIRST MOVES, AND SUBJECT GROUPS

Classification			Sources ^a						
Task	First move	Subjects	Human <i>vs.</i> machine (1)	Sex (2)	Task order (3)	Inter- action (1 x 2)	Inter- action (1 x 3)	Inter- action (2 x 3)	Inter- action (1 x 2 x 3)
I	Experimenter	A	2.10	1.74	2.94	1.74	3.91	1.74	1.74
I	Experimenter	B	2.59	9.07**	1.77	.63	2.59	.04	1.42
I	Subject	A	.01	10.43**	4.52*	3.28	.31	1.40	1.05
I	Subject	B	.17	5.17*	4.66*	.17	3.27	.09	.03
II	Experimenter	A	.73	.00	2.25	.18	.41	2.25	.04
II	Experimenter	B	.32	11.07**	.32	.01	1.90	.01	.53
II	Subject	A	.10	1.45	.33	.00	.33	.10	4.92*
II	Subject	B	.67	4.42	.31	1.34	.53	1.15	.67

^a All ratios use individual difference chi square as the "error" term.

* Significant at the .05 level.

** Significant at the .01 level.

transfer of training provided by experience with Task II when it preceded Task I. The interesting observation is that such transfer did *not* occur at a significant level for the remaining six categories. The F ratio for none of the six reached the .10 level of significance. No ready explanation for this phenomenon seems to be available.

The most significant finding shown by results summarized in Table 3 is that none of the F ratios for condition (Column 1) is significant at the .10 level of confidence. There is no evidence suggesting that the subjects performed better under either the "machine" or the "human" sets. Thus, it would appear that effectiveness on the task as measured by the criterion of number of moves remaining does not depend upon the difference between a set to work with another human as a partner or a set to work with a machine as a partner.

E. PREDICTIONS

A part of the task assigned to each subject was predicting the move about to be made by either the "machine" or "partner" with whom he was working. This feature of the experiment was designed to yield information about what moves the subjects came to expect as the responses to their own acts. Since the subject made predictions only on alternate moves (i.e., before the experimenter moved), only the first three predictions per trial were completely free of the ambiguity that could be produced by the fact that the task had been completed in the minimum number of moves. Only these latter data, or a total of 48 predictions per subject, were retained for analysis. A comparison was made between the number of trials for which perfect predictions and perfect performance were achieved by each subject.

Perfect prediction was ascertained for three criteria of perfection. The first (Criterion A) required that each of the three predictions be exactly the square that the experimenter selected on the predicted move. The second (Criterion B) required only that the predicted square be one that could lead to a perfect solution of the problem if the pattern in the experimenter's moves was not considered as a factor. This criterion of accurate prediction required an awareness on the part of the subject of possible "block" in future moves, but did not require that he correctly indicate the exact square to which the experimenter subsequently moved. The third (Criterion C) was the most relaxed in that the subject was required only to observe the rules of determining legal moves and to note that the predicted move was a forward move (i.e., would reduce the number of moves remaining).

Inspection of the data showed most clearly that many of the subjects were

successful in completing the tasks in the minimum number of moves more frequently than they were successful in predicting correctly the first three moves of their partner. If we consider the data for Criterion A, only two of the 48 subjects earned a larger number of perfect predictions than of successful goal achievements and, in each of these two cases, the excess was a single success in prediction. In contrast, one subject achieved the goal successfully on 11 trials (62 per cent), but predicted correctly on only two occasions (12 per cent). And a second succeeded on eight trials (50 per cent), but predicted correctly only once (six per cent). A simple sign test (9) shows these differences between the frequency of perfect prediction and of perfect goal achievements to be highly significant ($p < .001$).

The Criterion B data changed the picture somewhat. Twenty-five subjects made a larger number of correct predictions than successful goal achievements; however, detailed examination of data showed that 42 subjects succeeded in reaching the goal one or more times on trials for which they had made incorrect predictions. Using the most relaxed criterion, C, for which 47 of the subjects predicted correctly more frequently than they were successful in goal achievement, detailed data showed that 20 of the subjects obtained the goal in the minimum number of moves on one or more trials on which they had made an error in prediction; i.e., had not predicted a "forward" move for the experimenter.

The conclusion that must be drawn is that some subjects, at least part of the time, were able to complete the tasks successfully without being capable of predicting what response would be made to their moves. It appears that there must be problem situations (tasks) that have the basic characteristic of a mutual problem, but within which an individual is able to work effectively without first learning to predict how the other will respond to his acts. One possible means of ascertaining whether task effectiveness precedes accuracy in prediction is to construct mutual tasks with specific additional characteristics. These characteristics may furnish the key to the disparity.

In order to test the possible effects of the experimental sets of "human" *vs.* "machine" partners upon the prediction accuracy of the subjects, analyses were made separately of the frequency data for Subjects A and for Subjects B. The rationale for this division has been suggested previously.

This analysis involved a simple partition of χ^2 into components associated with the three major factors in the experimental design. The basic data were the number of correct predictions made by each subject on the first three moves of the "other" for each trial summed over 16 trials.

The only significant factor related to accuracy of prediction found for the

subjects in Group A was the sex of the subject. The 12 women subjects achieved an average accuracy of prediction of 25.0 correct, 52 per cent of the 48 predictions that each made. The 12 men subjects averaged 32.3 accurate predictions in 48 opportunities, or approximately 67 per cent. The "human" *vs.* "machine" sets given the subject were not factors affecting the accuracy of prediction.

For the 24 subjects in Group B, the difference between sexes in the frequency with which correct predictions of the move of the "other" were made is again significant ($p < .02$). The 12 men subjects predicted accurately 64 per cent of the time as compared with 52 per cent for the women subjects.

F. PATTERNS OF ACTS AND RESPONSES

It is possible for the work of the subjects on the tasks to differ with the conditions of the experiment in ways that would not be reflected in the effectiveness of task performance, e.g., different patterns of responses associated with preferences for certain routes or general approaches to the goal rather than others.

In order to make the analysis of patterning effect practical, it was necessary to establish general "indicators" of major pattern differences. Three such indicators were constructed for Task I and three for Task II. The first step in constructing these indices was to plot actual moves made by each subject on each trial. The four indices obtained are as follows:

1. The number of moves that were made to squares in Column D.
2. The number of moves that were made to squares in Row 2.
3. The number of moves made to squares in Column B (Task I only).
4. The number of moves made to squares in Column C (Task II only).

Counts of frequencies for these indices were limited to the first seven or the first eight moves depending upon the task. Chi-square analyses were made on data generated by application of the four indices to the records of the work of 24 subjects in Group A. Although some meaningful differences in approach to the tasks were found, the most important finding of the analysis for the present investigation is the absence of significant F ratio for the effects of the different sets. There is no reason to assume that these sets caused the subjects to approach the tasks in different manners.

G. DISCUSSION

The findings of this investigation of the effects of a set to work with a "machine" partner *vs.* a "human" partner in mutual problem solving have been generally negative. It was shown that the "sets" produced meaningful

differences in the attitudes of the subjects toward the way in which the partner was responding to the problem. "Human" partners tended to be regarded as more confused in their work than were "machine" partners. This situation can be understood best as a projection of the subject's own confusion, since objectively there was little or no difference between the actual responses of their partners. "Humans" easily can be regarded as confused, but such a characterization of "machines" does not seem appropriate. However, regardless of this attitudinal difference produced by the sets, the actual problem-solving behavior of the subject was not significantly affected.

Measured by (a) accuracy of predicting the response of the partner, (b) effectiveness of work on the task, and (c) patterns of approach to the task, a null hypothesis of no difference between sets would not have been refuted. The fact that each of these measures was sufficiently sensitive to show sex difference, task-order difference, or difference associated with the interaction of variables gives added confidence to the view that being set to work with a "human" partner, as compared with a "machine" partner, was not a potent source of variance in the subjects' problem-solving behavior on these tasks.

It would not be appropriate to generalize the finding of this study to tasks that differ markedly from those used in the experiment. These tasks were deliberately chosen to be simple; no communication other than that contained in the signal character of the partner response was permitted, and subject matter or content of the tasks had a definite "game" character. It seems very likely that tasks could be found for which significant difference with "human" and "machine" sets could be demonstrated.

The significance of the present result lies in the suggestion that there are important task properties that are common to man-with-man systems and man-with-machine systems. In current applied psychological work, there is much concern with man-machine systems. Man's long concern with problems of the effectiveness of his working relationships with his fellow men is not in need of documentation. It appears that effective relationships between two interacting entities may have significant characteristics that can be studied at a most general level. Understanding of the complex problems of social interactions may be facilitated by first developing a better knowledge of the strictly nonsocial elements that social interactions embrace.

H. SUMMARY AND CONCLUSIONS

An experiment was designed to explore differences that might exist between groups involved in mutual problem solving when they are (a) "set" to work with a human partner or (b) "set" to work with a machine as a partner.

Participating in the experiment were 24 men and 24 women; half of each sex were run under each experimental condition.

The problem involved the use of a plugboard containing 25 squares arranged in a five-by-five configuration, with one square designated as the goal square. The object was to achieve the goal in as few moves as possible with each subject alternating moves with his "partner" or the "machine." The perfect solution in Task I required eight moves and in Task II, seven moves.

Three measures of performance were taken: (a) number of moves remaining to be made after the minimum required had been exhausted, (b) patterns that developed in moves toward the goal square, and (c) accuracy of prediction of the "partner's" next move. The moves-remaining score showed no significant differences that were associated with the contrasting "human" *vs.* "machine" sets. The three indices of patterning of moves showed that the subjects with the "human" sets did not use routes to the goal that differed from those used by subjects with the "machine" sets. The subjects participating in one condition were not better able to predict the next move of their "partner" than those subjects participating in the other.

An attitude questionnaire did reveal that those subjects who had been given the "human" set considered their partners as more confused than did those working under the "machine" set.

Although the results of this experiment cannot be generalized to all tasks, they suggest the existence of important structural components in certain tasks that are not dependent upon the social character of group problem solving. Much further study of these components will probably be required before we can attain a proper understanding of their significance in relation to social and personal variables that are also involved in group problem solving.

REFERENCES

1. ANDERSON, H. Group performance in an anagram task. *J. Soc. Psychol.*, 1961, **55**, 67-75.
2. ANDERSON, T. W. *Introduction to Multivariate Statistical Analysis*. New York: Wiley, 1958.
3. HEMPHILL, J. K. Administration as problem-solving. In A. W. Halpin (Ed.), *Administrative Theory in Education*. Chicago: Midwest Administration Center, 1958.
4. ———. The function of structure-in-interaction in mutual problem solving. ONR technical report, Educational Testing Service, Princeton, N. J., February, 1961.
5. MATHER, K. *Statistical Analysis in Biology*. London: Methuen, 1943.
6. McNEMAR, Q. *Psychological Statistics*. New York: Wiley, 1962.
7. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957.

8. PEPINSKY, P. N., HEMPHILL, J. K., & SHEVITZ, R. N. Attempts to lead, group productivity, and morale under conditions of acceptance and rejection. *J. Abn. & Soc. Psychol.*, 1958, **57**, 47-54.
9. SIEGEL, S. *Nonparametric Statistics*. New York: McGraw-Hill, 1956.
10. THIBAUT, J. W., & KELLY, H. H. *The Social Psychology of Groups*. New York: Wiley, 1959.
11. WINER, B. J. *Statistical Principles in Experimental Design*. New York: McGraw-Hill, 1962.

Developmental Research Division
Educational Testing Service
Princeton, New Jersey

CHARACTERISTICS OF NURSES IN SHORT- AND LONG-TERM TREATMENT SETTINGS*

Department of Psychiatry, D. C. General Hospital

JOAN K. BORUCHOW^{1, 2}

A. INTRODUCTION

As the service professions (such as medicine, nursing, and social work) have become increasingly specialized, the traditional stereotype of members of these professions as warm, giving, interested in people, and so forth has perhaps less relevance than was the case 10 or 20 years ago. A comparative study of neuropsychiatric and general medical-surgical nurses by Navran and Stauffacher (4) is illustrative of recent efforts to identify personality characteristics associated with subspecialties in these professions. They found general medical-surgical nurses to be significantly more orderly, deferent, and self-abasing and less introspective and aggressive than neuropsychiatric nurses. The interpretation was that the general medical-surgical nurses are more work oriented, impersonal, and less able to direct others than are their psychiatric colleagues, who place greater emphasis on interpersonal relationships. However, there is some difficulty in interpreting these results as the authors do not specify whether long-term and short-term treatment settings were equally represented in their samples. Consequently, the differences may reflect differences in intensity and duration of patient contact irrespective of the type of nurse. In the present study an effort was made to compare and contrast characteristics associated with nurses in long-term and short-term treatment settings and characteristics associated with nurses in neuropsychiatric and general medical-surgical settings.

B. METHOD

1. *Subjects*

Subjects were 160 registered nurses representative of four types of treatment settings. Each setting included 40 nurses. Psychiatric units in receiving

* Received in the Editorial Office, Provincetown, Massachusetts, on March 30, 1964. Copyright, 1965, by The Journal Press.

¹ This study was supported by funds made available by Research Grant MHO 4663-03 of the Psychopharmacology Service Center, National Institute of Health, U. S. Public Health Service.

² The author wishes to thank Dr. Allen Raskin for his generous assistance.

hospitals³ and state mental hospitals⁴ represented the short-term and long-term settings in psychiatry respectively. The operating room⁵ represented the short-term general medical-surgical setting.

As medical-surgical treatment of tuberculosis⁶ and orthopedic⁷ patients involves prolonged care and long contact, the nurses from these areas were combined to represent the long-term general medical-surgical setting. The orthopedic and tuberculosis groups were compared on background and personality variables to see if they were homogeneous and could be combined. These tests revealed significant differences on three of the background variables and one of the personality variables. Consequently, for these four variables, the orthopedic and tuberculosis groups were *not* combined in testing differences among the various nursing specialties.

The nine background and 42 personality variables were incorporated into a five-part, self-administering questionnaire in booklet form. The forms were mailed to the director of nursing of each hospital for distribution to the staff nurses. To assure anonymity the booklet listed a code number to identify the subject and setting, with a request to omit the subject's name. To discourage the tendency to give socially desirable responses, the subject was instructed to mail the forms directly to the author instead of returning them to the director of nursing.

2. Measures

a. Background variables. These were age, race, marital status (married or other), education, earnings, intent to remain, length of employment, and contentment with present position.

b. Personality variables. These were as follows: (a) somatic complaints—16 items describing common physical complaints, (b) psychic complaints—24 items descriptive of emotional complaints [Variables (a) and (b) were adapted from Frank's Symptom Distress Scale], (c) sociability—15 items from the Sociability Scale of the Guilford-Zimmerman Temperament Survey, (d) democratic attitude—25 items adapted from the Adorno-Levinson F Scale,

³ Crownsville State Hospital, Crownsville, Md.; Malcolm Bliss Mental Health Center, St. Louis, Mo.; Spring Grove Hospital, Crownsville, Md.; Springfield Hospital, Springfield, Md.

⁴ Crownsville State Hospital, Crownsville, Md.; Eastern Shore State Hospital, Cambridge, Md.; Kentucky State Hospital, Danville, Ky.; Saint Elizabeths Hospital, Washington, D. C.

⁵ George Washington Hospital, Washington, D. C.; Georgetown Hospital, Washington, D. C.; Providence Hospital, Washington, D. C.; Washington Hospital Center, Washington, D. C.

⁶ National Orthopedic and Rehabilitation Hospital, Arlington, Va.; North Carolina Orthopedic Hospital, Gastonia, N. C.; Suburban Hospital, Arlington, Va.

⁷ Glenn Dale Hospital, Glenndale, Md.

(*e*) subservience to M.D., (*f*) egalitarianism, (*g*) impersonal attitude, (*h*) task orientation, (*i*) status distance, (*j*) belief in personal efficacy, (*k*) belief in efficacy of nonmedical personnel, (*l*) personal contact and involvement, (*m*) belief in order and control [Variables (*e*) through (*m*) included from four to seven items and were adapted from the National Institutes of Health Nursing Attitude Survey], (*n*) autocratic, (*o*) distrustful, (*p*) overconventional, (*q*) exploitive, (*r*) self-effacing, (*s*) hostile-aggressive, (*t*) dependent, (*u*) overgenerous, (*v*) managerial, (*w*) cooperative, (*x*) self-reliant, (*y*) forthright, (*z*) accepting, (*aa*) responsible [Variables (*n*) through (*aa*) included from four to 12 items adapted from the Leary Interpersonal Check List (3); Variables (*n*) through (*u*) included positive statements, and variables (*v*) through (*aa*) were negative], and (*bb*) self-ideal nurse agreement [the 128-item Leary Interpersonal Check List was rated for the "Self" and again for the "Ideal Nurse"]. Phi coefficients used as a measure of agreement.

Analyses of variance and chi-square tests were computed to test for differences among the four settings on the background variables. Covariance analyses, controlling for the tendency to give socially desirable responses (2), were computed to test for differences among the four treatment settings on the personality variables. When there was an overall significant difference between groups, Duncan's multiple-range test was employed to determine where the groups differed.

C. RESULTS AND DISCUSSION

The two general medical-surgical groups differed from the two psychiatric groups on only one variable: egalitarianism. This variable included items reflecting class and status distinction among hospital-staff members and among patients and hospital staff. Psychiatric nurses scored higher than the general medical-surgical nurses, indicating a more democratic attitude.

Nurses in short-term and long-term settings also differed significantly on only one variable; nurses from the short-term settings perceived the "ideal nurse" as more dependent than did nurses in long-term settings.

In contrast to the minor differences between nurses in psychiatric and general medical-surgical settings and between nurses in long-term and short-term treatment settings, there were seven significant differences between the groups representing the extremes in intensity and duration of patient contact: i.e., the operating-room nurses and nurses from state mental hospitals. Operating-room nurses rated themselves as more dependent than did the other three groups, a factor that can be related to the background variables that showed [see Table 1] that although they are the lowest salaried and indicated least

intention to remain in their present position, they have remained in their position on the average of 34 months.

The personality variables (listed in Table 2) that can be loosely interpreted as interpersonal-relationship measures show that operating-room nurses perceived themselves as more introverted and place less importance than others

TABLE 1
BACKGROUND VARIABLES THAT SIGNIFICANTLY DIFFERENTIATED AMONG
THE FOUR NURSING GROUPS

Variable	Neuropsychiatric nurses		General medical-surgical nurses		F
	Short term	Long term	Short term	Long term	
Age (in years)	30	35	32	44	10.68**
Race (per cent white)	33	88	98	98	69.77***
Married (per cent)	67	70	55	52	3.51*
Salary (in dollars)	\$4450	\$4350	\$3820	\$4999	3.83*
Supervisory nurses (per cent)	50	68	20	48	14.25**
Intention to remain (in months)	23	24	15		3.41*
Months in position	26	30	34	55	2.80*

* Significant at the .05 level.

** Significant at the .01 level.

^a Chi square.

TABLE 2
MEAN SCORES OF PERSONALITY VARIABLES THAT SIGNIFICANTLY DIFFERENTIATED
AMONG THE FOUR NURSING GROUPS

Variable	Neuropsychiatric nurses		General medical-surgical nurses		F
	Short term	Long term	Short term	Long term	
Subservience to M.D.	11.70	10.90	12.60	11.42	7.94**
Egalitarianism	17.50	17.48	16.20	16.19	5.51**
Belief in efficacy of nonmedical personnel	12.90	12.98	12.28	11.28	6.07**
Autocratic (self)	1.10	1.15	.87	.35	4.09**
Overconventional (self)	1.81	1.34	2.31	1.72	2.93*
Self-effacing (self)	3.14	2.47	4.75	3.89	8.35**
Hostile (self)	1.90	1.94	2.86	1.69	3.78*
Dependent (self)	3.85	3.23	5.54	4.33	10.12**
Accepting (ideal)	3.44	3.05	3.65	3.29	3.26*
Autocratic (ideal)	1.74	1.17	1.37	.71	2.76*
Dependent (ideal)	3.56	2.60	3.45	2.84	3.45*
Self-ideal nurse agreement (phi)	.62	.69	.53	.56	8.76**
Democratic attitude	69.64	65.34	61.96	64.25	4.42**

* Significant at the .05 level.

** Significant at the .01 level.

on interpersonal relationships. Their ratings were highest on the personality variables, "overconventional" and "hostile," suggesting less freedom and satisfaction in their interpersonal relationships. In addition, they saw themselves as more self-effacing than did the other three groups and were lowest on self-ideal nurse agreement. The latter result is consistent with Raskin's findings on psychiatric outpatients, who rated themselves as less self-satisfied, sociable, and assertive and had poorer self admired-other agreement scores (5). Operating-room nurses were also lowest in democratic attitude, indicating greater need for direction from authority figures and less desire for self-expression and authority in their own positions. This finding relates also to their high subservience-to-M.D. scores.

Long-term psychiatric nurses were opposite to their operating-room colleagues on dependency and subservience to M.D. They also placed more value on democratic thought (interpreted as less attachment to authority and more on self-expression), with much interest in interpersonal relationships. They were highest on self-ideal nurse agreement, presenting a higher self-concept and perhaps indicating more satisfaction with their chosen profession (1).

There were also a few significant characteristics associated with nurses in tuberculosis and orthopedic settings. The tuberculosis nurses were older, had been employed in their present positions longest (mean of 67 months), and intended to remain in their positions longer than did the typical member of each of the other groups. Orthopedic nurses differed from nurses in other settings on only one variable; they were lowest in belief in efficacy of non-medical personnel. However, considering the number of tests of significance, caution should be exercised in interpreting these differences as they could occur by chance alone.

In sum, there were many characteristic differences between nurses on the extremes of a continuum representing intensity and duration of interpersonal contact with patients: i.e., the operating-room and the long-term psychiatric nurses. Further, these differences were in the direction predicted by the Navran and Stauffacher (4) study. However, there were no differences between groups representing less extreme positions on this continuum (i.e., between short-term psychiatric nurses who have relatively intense interpersonal contact with patients for short periods of time and long-term general medical-surgical nurses who have less intense interpersonal contacts with patients for longer periods of time). In other words, categorizing nursing settings on the basis of both intensity and duration of nurse-patient contact apparently provides a more meaningful distinction in terms of personality attributes of nurses than

the distinction between psychiatric and general medical-surgical nurses per se. In this regard it would be of interest to know if comparable results would obtain if members of other service professions were similarly categorized on the basis of intensity and duration of patient contact. For example, do Travelers Aid Society social workers, who usually have brief and limited contact with clients, differ from social workers in state mental hospitals? and are these differences in personality attributes in the direction predicted from the results of the present study?

D. SUMMARY

Background and personality characteristics of 160 psychiatric and general medical-surgical nurses in short-term and long-term treatment settings were compared and contrasted. There were no marked differences between general medical-surgical and psychiatric nurses generally or between nurses in short-term and long-term settings. The major differences were between the two extreme groups on the dimension of intensity and duration of patient contact: i.e., operating room and state-hospital psychiatric nurses. As compared to state-hospital psychiatric nurses, operating-room nurses were more dependent, submissive to the physician, had lower self-concepts and preferred being directed rather than assuming authority themselves. In contrast, psychiatric nurses placed more value on interpersonal relationships and were more outspoken and extroverted than were their operating-room colleagues.

REFERENCES

1. BROPHY, A. L. Self, role and satisfaction. *Genet. Psychol. Monog.*, 1959, **59**, 263-308.
2. CROWNE, D. P., & MARLOW, D. A. A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 1960, **24**, 349-354.
3. LAFORGE, R., & SUCZEK, R. F. The Interpersonal Dimension of Personality: An Interpersonal Checklist. *J. Personal.*, 1955, **24**, 94-112.
4. NAVRAN, L., & STAUFFACHER, J. C. A comparative analysis of the personality structure of psychiatric and non-psychiatric nurses. *Nursing Res.*, 1958, **7**, 64-67.
5. RASKIN, A. Factors that influence or are associated with self-admired other concomitance indices of psychotherapy outpatients. *J. Clin. Psychol.*, 1962, **18**, 37-39.

Research Department, Psychiatry
Mercy Douglass Hospital
5000 Woodland Ave.
Philadelphia, Pa. 19143

SOME PERSONALITY DIFFERENCES BETWEEN DOGMATIC AND NONDOGMATIC GROUPS*¹

Department of Psychology, San Jose State College

WALTER T. PLANT, CHARLES W. TELFORD, AND JOSEPH A. THOMAS

A. INTRODUCTION

The purpose of this paper is to present the results of analyses directed toward determining how persons who differ markedly in measured dogmatism differ in certain other respects. The Rokeach Dogmatism Scale: Form E was the measure used to select extreme groups for this purpose.

In the monograph *Political and Religious Dogmatism* (13), Rokeach published a 40-item measure of general authoritarianism or dogmatism. Since its publication, the Dogmatism Scale (hereinafter referred to as the D Scale) has been and continues to be used extensively in social-psychological research and particularly in research dealing with attempts to measure the non-intellectual effects of a college education (4, 8, 12, 15). The most comprehensive summary of work with the scale to date is found in Rokeach's book *The Open and Closed Mind* (14).

In the development of the D Scale, systematic effort demonstrated that this scale is less loaded with a left-right or liberal-conservative political dimension than is the familiar California F Scale, and it contains less avowed-prejudice content than does the California Ethnocentrism Scale (or E Scale). Intercorrelational studies of the D Scale and F Scale with measures of political and economic conservatism have consistently yielded lower correlations for the D Scale than for the F Scale. An examination of scale content yields the observation that the D Scale contains much less avowed-prejudicial material or stated-prejudicial material than does the E Scale. Rokeach has also reported intercorrelational studies with the D, E, and F Scales from which he has concluded that the D Scale is a better measure of general authoritarianism than is the F Scale. Replications of certain of these correlational studies (11) with American subjects, as against Rokeach's English subjects, have reported general agreement with the Rokeach findings. Rokeach (13) has reported odd-even split-half reliability coefficients of

* Received in the Editorial Office, Provincetown, Massachusetts, on April 1, 1964.
Copyright, 1965, by The Journal Press.

¹ The data for this study were obtained as a part of the research program under Contract 914, U. S. Office of Education, Cooperative Research Branch (15).

.81 for a small sample of English college students and .78 for a sample of English workers. Plant (12) has reported odd-even split-half reliability coefficients of .84 for a sample of 400 male college freshmen and .85 for a sample of 400 female college freshmen. The scale is sufficiently reliable for use in research dealing with group differences.

Peabody (10) has reported extensive work on response set and the D Scale. He used the D Scale, the F Scale, an anti-Semitism scale, a conservatism scale, and reversed scales for each. Peabody, using both English and American samples of subjects, has concluded that low scores on the D Scale represent rejection of dogmatism; whereas high scores represent a combination of "agreement set" or acquiescence and acceptance of dogmatism.

There have been studies aimed at determining how groups of subjects with extreme scores on the D Scale differ also in other respects. Rokeach (14) demonstrated that persons with relatively high D-Scale scores solve complex laboratory problems much less well than do persons with relatively low D-Scale scores. The problems that have been used require that a subject be able to overcome sets of beliefs about the nature of things that ordinarily are used in complex-problem solution and to adopt new approaches to the problem and its solution. Fillenbaum and Jackman (3) have replicated the Rokeach studies with subjects having high and low D-Scale scores using the same complex-problem task and have obtained essentially the same results. In addition, these investigators have demonstrated that subjects with low scores on the D Scale also have relatively low scores on a measure of anxiety and that those with high scores on the D Scale have high scores on the anxiety measure.

It would seem that the D Scale is a reasonably good measure of general authoritarianism or dogmatism and is one for which some information is available about subjects who score relatively high or low on it. Our intent is to add to the information on subjects who are high and those who are low in dogmatism as measured by the D Scale.

B. MEASURES USED AND METHODS OF ANALYSIS

1. *Measures Used*

In the summer and fall of 1960, a battery of psychological tests was given to all persons indicating an intent to enroll for the academic year 1960-1961 as freshmen in one of six California public junior colleges (15). The total number of subjects for whom usable test results were obtained was 4,506 (2,643 males and 1,863 females). All subjects were given the special battery, a scholastic-aptitude measure, and various achievement measures used in the

several colleges for placement purposes. The special battery included the following group measures:

1. The Rokeach Dogmatism Scale: Form E (13).
2. The Modified California Psychological Inventory (5), hereinafter referred to as the CPI. The complete Gough inventory includes 18 scales. The modified inventory includes five of these scales taken intact and printed into a new booklet. The five scales, chosen on the basis of local factor-analysis studies (6), were the sociability (Sy), self-control (Sc), achievement via independence (Ai), intellectual efficiency (Ie), and responsibility (Re) scales.
3. The Allport, Vernon, and Lindzey Study of Values (1) hereinafter referred to as the AVL. This widely known scale yields scores in the following value or interest domains: theoretical, economic, aesthetic, social, political, and religious.

The scholastic-aptitude measure was the School and College Ability Test (SCAT): Form A, or another similar measure, the scores for which were converted to SCAT total raw scores by a conversion procedure offered by the Psychological Corporation (2).

2. *Methods of Analysis*

Two sets of statistical analyses were made of the data from the test battery for extreme groups on the D Scale. For the analyses of Study I, the top and bottom 10 per cent of the scores on the D Scale were selected from the sample of 2,643 males and from the sample of 1,863 females. Means and *t* ratios for the five CPI-scale scores, the six AVL-scale scores, and the scholastic-aptitude-scale score were computed for the high D-Scale score males *vs.* the low D-Scale score males and for the high D-Scale score females *vs.* the low D-Scale score females. Separate analyses for males and for females were undertaken because of the repeated finding of significant sex differences on the D Scale (8, 12, 15), on the CPI scales (5), and on the AVL scales (1).

For the analyses of Study II, 110 pairs of males and 110 pairs of females were selected for study. The members of the pairs were matched within one raw-score point on the total raw score from the SCAT measure, but one member of each pair had a score from the highest 15 per cent on the D Scale while the other member of the pair had a score from the lowest 15 per cent on the D Scale. Statistical tests of the significance of the difference between correlated means were undertaken for the 110 matched pairs of males and for the 110 matched pairs of females on the five CPI-scale and the six AVL-scale scores.

C. RESULTS AND DISCUSSION

1. Study I

The results of the tests of the significance of the difference between sample means for the five CPI scales and the SCAT total raw scores for low D-Scale males *vs.* the high D-Scale males are given in Table 1. The same kind of results for the same variables for females with low D-Scale scores *vs.* females with high D-Scale scores are given in Table 2.

Each of the comparisons between dogmatic males and nondogmatic males reported in Table 1 yields a mean difference significant beyond the .01 level.

TABLE 1
CPI AND SCAT COMPARISONS BETWEEN MALES LOW IN MEASURED DOGMATISM
AND MALES HIGH IN MEASURED DOGMATISM

Variable	Low dogmatism (<i>N</i> = 264)		High dogmatism (<i>N</i> = 264)		Difference	<i>t</i>
	Mean	<i>SD</i>	Mean	<i>SD</i>		
D Scale, Form E	119.3	12.5	213.9	11.8		
California Psychological Inventory						
Sociability (Sy)	25.1	5.1	22.2	4.7	2.9	6.88*
Self-control (Sc)	28.4	8.2	21.6	7.7	6.8	9.86*
Achievement via independence (Ai)	20.4	4.6	12.8	3.8	7.6	20.68*
Intellectual efficiency (Ie)	37.9	6.2	30.1	6.3	7.8	14.40*
Responsibility (Re)	29.1	5.4	25.8	5.2	3.3	7.16*
School and College Ability Test: Total	68.1	18.7	52.1	15.1	16.0	10.78*

* $p < .01$.

TABLE 2
CPI AND SCAT COMPARISONS BETWEEN FEMALES LOW IN MEASURED DOGMATISM
AND FEMALES HIGH IN MEASURED DOGMATISM

Variable	Low dogmatism (<i>N</i> = 186)		High dogmatism (<i>N</i> = 186)		Difference	<i>t</i>
	Mean	<i>SD</i>	Mean	<i>SD</i>		
D Scale, Form E	111.1	11.8	211.2	12.3		
California Psychological Inventory						
Sociability (Sy)	25.1	5.5	21.6	6.7	3.5	5.50*
Self-control (Sc)	30.2	8.1	24.8	8.7	5.4	6.17*
Achievement via independence (Ai)	21.2	3.9	14.0	4.0	7.2	17.36*
Intellectual efficiency (Ie)	38.5	5.8	31.0	6.4	7.5	11.87*
Responsibility (Re)	32.1	4.6	28.5	5.4	3.6	6.83*
School and College Ability Test: Total	62.8	19.1	47.8	16.7	15.0	8.07*

* $p < .01$.

Without exception, the nondogmatic males had the highest means on the five CPI variables and on the SCAT variable.

Each of the comparisons between dogmatic females and nondogmatic females reported in Table 2 yields a mean difference significant beyond the .01 level. Without exception, the nondogmatic females had the highest means on the five CPI variables and on the SCAT variable.

The results of the tests of the significance of the difference between sample means for the six AVL scales yield one statistically significant difference for males and one for females. The nondogmatic males had a significantly lower mean on the AVL religious scale than did the dogmatic males. The nondogmatic females had a significantly higher mean on the AVL aesthetic scale than did the dogmatic females. In these two instances of statistical significance, the score differences between sample means are small relative to the average of the two standard deviations for the means. In addition, when controls on ability differences were made for the analyses of Study II, none of the AVL differences between means for dogmatic males *vs.* nondogmatic males or for dogmatic females *vs.* nondogmatic females was significant. In light of these considerations, no empirical results for the AVL comparisons will be reported.

2. Study II

The analyses made for Study I were repeated for samples of subjects high and low on the D Scale but matched on ability-test score because in Study I the differences between high and low dogmatics on the SCAT total raw score were substantial. In addition, studies of very bright *vs.* very dull subjects (9) report significant differences on CPI variables, with the very bright subjects having the highest means; and studies of very bright subjects (7, 16) report relatively high mean scores on the CPI scales. The purpose of these analyses was to determine if there were personality differences between those high and those low in measured dogmatism but alike in measured scholastic aptitude.

The results of the tests of the significance of the difference between correlated sample means for the five CPI scales for low D-Scale males *vs.* high D-Scale males but matched on SCAT scores are found in Table 3. The same kind of results for the same variables for females with low D-Scale scores *vs.* females with high D-Scale scores but matched on SCAT scores are found in Table 4.

Each of the comparisons reported in Table 3 between dogmatic males and nondogmatic males who were matched on ability-test scores yields a mean

TABLE 3
CPI COMPARISONS BETWEEN MALES LOW IN MEASURED DOGMATISM AND MALES
HIGH IN MEASURED DOGMATISM BUT MATCHED ON ABILITY

Variable	Low dogmatism (<i>N</i> = 110)		High dogmatism (<i>N</i> = 110)		Difference	<i>t</i>
	Mean	<i>SD</i>	Mean	<i>SD</i>		
D Scale, Form E	115.7	14.8	211.9	12.9		
California Psychological Inventory						
Sociability (Sy)	25.5	5.2	21.9	4.7	3.6	5.47*
Self-control (Sc)	29.7	7.5	21.7	7.8	8.0	7.72*
Achievement via independence (Ai)	20.7	4.2	14.0	3.9	6.7	14.44*
Intellectual efficiency (Ie)	38.7	4.9	31.9	6.7	6.8	9.62*
Responsibility (Re)	30.0	5.0	26.9	5.2	3.1	4.59*

* $p < .01$.

TABLE 4
CPI COMPARISONS BETWEEN FEMALES LOW IN MEASURED DOGMATISM AND FEMALES
HIGH IN MEASURED DOGMATISM BUT MATCHED ON ABILITY

Variable	Low dogmatism (<i>N</i> = 110)		High dogmatism (<i>N</i> = 110)		Difference	<i>t</i>
	Mean	<i>SD</i>	Mean	<i>SD</i>		
D Scale, Form E	116.2	13.0	204.4	14.0		
California Psychological Inventory						
Sociability (Sy)	24.6	5.7	23.0	7.2	1.6	1.99*
Self-control (Sc)	31.0	8.2	24.7	8.5	6.5	5.88**
Achievement via independence (Ai)	20.2	4.2	15.3	4.1	4.9	9.50**
Intellectual efficiency (Ie)	37.2	6.0	32.8	5.5	4.4	6.45**
Responsibility (Re)	31.5	4.8	29.3	6.0	2.2	3.29**

* $p < .05$.

** $p < .01$.

difference significant beyond the .01 level. Without exception, the nondogmatic males had the highest means on the five CPI variables.

Four of the comparisons reported in Table 4 between dogmatic and nondogmatic females who were matched on ability-test scores yields a mean difference significant beyond the .01 level. One comparison, on the sociability scale (Sy) yields a mean difference significant beyond the .05 level. Without exception, the nondogmatic females had the highest means on the five CPI variables.

None of the tests of the significance of the difference between sample correlated means for high D-Scale males *vs.* low D-Scale males or for high D-Scale females *vs.* low D-Scale females on the AVL was statistically significant.

3. Discussion

When comparisons were made of measured personality characteristics of persons with extreme scores on the D Scale, statistically significant differences were obtained on the five scales of the CPI and on the SCAT. This was true for samples of males and of females, and (without exception) the relatively nondogmatic subjects had the highest mean scores on the CPI and SCAT scales. Compared to highly dogmatic subjects, the nondogmatic subjects would be described as being outgoing and enterprising, calm and patient, mature and forceful, efficient and clear thinking, planful and responsible, and more likely to succeed in an academic setting than would the highly dogmatic subjects. These descriptive adjectives come from the descriptions of the CPI scales and the extensive validity studies conducted by Gough (5). The description of the nonintellectual characteristics of those relatively free of dogmatism, compared with those high in dogmatism but alike in scholastic aptitude, is the same as reported earlier. Controls on the aptitude variable through the matched-pair procedure did not change the results.

Although all the differences between the low dogmatics and high dogmatics on the CPI variables were statistically significant, the differences on the achievement-via-independence and intellectual-efficiency scales were most striking. The mean differences on the first of the two aforementioned scales, without exception, were larger than the average of the two standard deviations of the means from which the differences were obtained. On three of the four comparisons reported for the intellectual-efficiency scale, the same kind of difference relative to the standard deviations was obtained.

A different type of comparison is provided between the CPI means obtained in this study and those obtained from Gough's normative samples (5). In Tables 5 and 6, the means for the high-school subjects and for the college

TABLE 5
CPI MEANS FOR MALE NORMATIVE SAMPLES AND LOW-DOGMATISM SAMPLE
AND HIGH-DOGMATISM SAMPLE

CPI Scale	High-school students Gough (5) (<i>N</i> = 3572)	College students Gough (5) (<i>N</i> = 680)	Low dogmatism (<i>N</i> = 110)	High dogmatism (<i>N</i> = 110)
Sociability (Sy)	21.5	25.4	25.5	21.9
Self-control (Sc)	25.3	29.2	29.7	21.7
Achievement via independence (Ai)	14.6	21.6	20.7	14.0
Intellectual efficiency (Ie)	33.6	40.8	38.7	31.9
Responsibility (Re)	26.7	31.5	30.0	26.9

subjects in Gough's standardization samples are presented, as are the means on the CPI scales for the high dogmatics and low dogmatics of the present study matched on scholastic aptitude. Table 5 contains the relevant means for males; Table 6, for females.

TABLE 6
CPI MEANS FOR FEMALE NORMATIVE SAMPLES AND LOW-DOGMATISM SAMPLE
AND HIGH-DOGMATISM SAMPLE

CPI Scale	High-school students Gough (5) (<i>N</i> = 4056)	College students Gough (5) (<i>N</i> = 2120)	Low dogmatism (<i>N</i> = 110)	High dogmatism (<i>N</i> = 110)
Sociability (Sy)	21.4	26.0	24.6	23.0
Self-control (Sc)	27.6	30.8	31.0	24.7
Achievement via independence (Ai)	15.5	21.9	20.2	15.3
Intellectual efficiency (Ie)	34.4	41.4	37.2	32.8
Responsibility (Re)	30.0	33.3	31.5	29.3

The mean CPI scores for the junior-college freshman males who had very low scores on the D Scale were very similar to those of Gough's college-student sample. The means for the freshman males who had very high scores on the D Scale were very similar to those of Gough's high-school sample and, in some cases, were below the level of the high-school sample. These differences seem to indicate that males with high D-Scale scores are likely to be psychologically immature and, although college freshmen, not as mature as many high-school subjects. The same general type of interpretation can be made for the female freshman subjects who had very high scores on the D Scale. These comparisons indicate that the dogmatic college freshman would be described as being impulsive, defensive, and conventional and stereotyped in thinking.

D. SUMMARY AND CONCLUSIONS

A battery of psychological tests, including the Rokeach Dogmatism Scale (D Scale), five scales from the California Psychological Inventory (CPI), the *Study of Values* (AVL), and a scholastic-aptitude test (SCAT), was given to a large sample of entering junior-college freshmen.

In order to study and to describe some of the ways in which highly dogmatic persons differ from nondogmatic persons, extreme groups on the D Scale were chosen. Comparisons for these extreme groups were made on CPI, AVL, and SCAT scores. In addition, the comparisons on the CPI and AVL were made for samples who had extreme scores on the D Scale but who were matched on aptitude-test score.

Nondogmatic males compared with dogmatic males and nondogmatic females compared with dogmatic females were found to differ significantly on each of the five CPI scales. Without exception, the nondogmatic groups had the highest means on the CPI scales. It was concluded that highly dogmatic subjects were psychologically immature and could be characterized as being impulsive, defensive, and conventional and stereotyped in thinking.

REFERENCES

1. ALLPORT, G. W., VERNON, P. E., & LINDZEY, G. Manual, Study of Values: A Scale for Measuring the Dominant Interests in Personality (3rd ed.) Boston: Houghton, 1960.
2. BENNETT, G. K., BENNETT, M. G., WALLACE, W. L., & WESMAN, A. G. Manual, College Qualification Tests. New York: Psychological Corp., 1957.
3. FILLENBAUM, S., & JACKMAN, A. Dogmatism and anxiety in relation to problem solving: An extension of Rokeach's results. *J. Abn. & Soc. Psychol.*, 1961, **63**, 212-214.
4. FOSTER, J., STANEK, R., & KRASSOWSKI, W. The Impact of a Value Oriented University on Student Attitudes and Thinking. Santa Clara, Calif.: Univ. Santa Clara, 1961.
5. GOUGH, H. G. Manual, California Psychological Inventory. Palo Alto, Calif.: Consulting Psychologists Press, 1957.
6. HICKS, R. A. Factor analytic studies of the California Psychological Inventory. Unpublished Master's thesis, San Jose State College, California, 1960.
7. HOLLAND, J. L. The prediction of college grades from the California Psychological Inventory and the Scholastic Aptitude Test. *J. Educ. Psychol.*, 1959, **50**, 135-142.
8. LEHMANN, I. J. Changes in critical thinking, attitudes, and values from freshman to senior years. *J. Educ. Psychol.*, 1963, **54**, 305-315.
9. NIX, J. T. Non-intellectual characteristics of gifted and non-gifted college students. Unpublished Master's thesis, San Jose State College, California, 1961.
10. PEABODY, D. Attitude content and agreement set in scales of authoritarianism, dogmatism, anti-Semitism, and economic conservatism. *J. Abn. & Soc. Psychol.*, 1961, **63**, 1-11.
11. PLANT, W. T. Rokeach's dogmatism scale as a measure of general authoritarianism. *Psychol. Reports*, 1960, **6**, 164.
12. ———. Longitudinal changes in intolerance and authoritarianism for subjects differing in amount of college education over four years. *Genet. Psychol. Monog.* (in press).
13. ROKEACH, M. Political and religious dogmatism: an alternative to the authoritarian personality. *Psychol. Monog.*, 1956, **70**, No. 18 (Whole No. 425).
14. ———. The Open and Closed Mind. New York: Basic Bks, 1960.
15. TELFORD, C. W., & PLANT, W. T. The Psychological Impact of the Public Two-Year College on Certain Non-intellectual Functions. San Jose, Calif.: San Jose State College, 1963.
16. WARREN, J. R., & HEIST, P. A. Personality attributes of gifted college students. *Science*, 1960, **132**, 330-337.

Department of Psychology
 San Jose State College
 San Jose 14, California

THE DEFINITION AND MEASUREMENT OF NATIONAL MORALE AND MORALITY*¹

*Department of Psychology, University of Illinois,
and
Department of Psychology, Vanderbilt University*

RAYMOND B. CATTELL AND RICHARD L. GORSUCH²

A. INTRODUCTION

The investigation of culture patterns by anthropologists and historians has long proceeded by methods that may be illustrated at their best by the writings of Benedict (1), Mead (20, 21), Keyserling (18), Toynbee (24), and others. A radically new approach, involving quantifying of observations and basing theory on an explicit, mathematical model was proposed by Cattell in 1951 (10); and it has generated concepts in studies by his associates which show the fruitfulness of the approach (6, 12). This new methodology (*a*) has taken nations as its organisms for correlation entries, on the thesis that these alone have sufficient autonomy for that internal feedback which guarantees the organic consequences of culture practices and thus makes factors meaningful; (*b*) has taken a wide array of variables—psychological, historical, economic, sociological—on the thesis that causal lines do not respect academic departments; (*c*) has distinguished three panels of variables—syntality, structure, and population—believing that these are closely connected by a stated formula (5); and (*d*) has employed simple-structure rotated factor-analytic solutions to discover independent, functionally unitary dimensions of national syntality, on the belief that the basic hypotheses in this area are best generated from correlational observations in detachment from armchair prejudices.

Among the 12 or more independent dimensions initially appearing in Cattell's study following this methodology (3, 10) was one tentatively designated "Poor Cultural Integration and Morale *vs.* Good Internal Morality." It loaded high death rate from syphilis, many deaths from alcoholism, many deaths from typhoid, tuberculosis, and other diseases brought under control in well-integrated communities, presence of restrictive laws against di-

* Received in the Editorial Office, Provincetown, Massachusetts, on April 2, 1964.
Copyright, 1965, by The Journal Press.

¹ The authors thank Harry C. Triandis for his many helpful comments, especially concerning the interpretation of some of the factors.

² This investigation was carried out while the junior author was supported by a Predoctoral Fellowship from the National Institute of Mental Health, United States Public Health Service, at the University of Illinois.

voice, and some less interpretable, lower variables, such as many miles of railroad per person. Some further light was thrown upon it by the P-technique factor-analytic studies (correlating time series over 100 variables for the factors in the country's historical development) by Gibb (12) for Australia, Cattell (4) for Great Britain, and Cattell and Adelson (6) for the United States. These showed a pattern tentatively called "Cultural Integration *vs.* Slum Standards" which similarly opposed community conscience and integration to carelessness, squalor, crime, and high disease rate. Although the samples and variables are not so exactly carried over from study to study as in the foregoing set of researches, the work of Hofstaetter (16) and Thorndike (23)—correlation clusters only—confirms a similar organic aggregate of measurable variables covering a variety of crime manifestations, absence of civic conscience (indeed, positive evidence of civic corruption), and substitution of crude, rigid, external controls for individual conscience. From these several studies, it is hypothesized that the present study will replicate the factor and find the interpretation of it supported by variables not previously included. It is also expected that evidence relating to alternative hypotheses, such as its being a health-related or slum factor, will be forthcoming.

The interpretation includes morale of the community as well as its morality on essentially two counts: (a) the nature of many of the variables that show at the pole of high morality a relatively well-integrated culture (e.g., lack of slum conditions) in which the members are interested in doing their share to uphold community standards and (b) research on smaller groups which has found that groups with high morale have high morality and low morale groups have low morality: e.g., Hartshorne and May (15). Unfortunately, the present investigation is limited in the clarification of this aspect, due to the nature of the variables available.

The interpretation we are seeking to clarify concerning moral behavior of individuals is in the sense of practicing values agreed on by the group which also transcend individual satisfactions; it would be one of several influences contributing to group synergy. [As seen through the education-affluence factor (3), educational influences are only related in a limited way.] The greater part of its variance may well arise from differences in willingness to accept ethical standards, but it nevertheless would be suspected of influencing a wide range of social behavior, including an ultimate effect on what communities are ready to pay for social services, as shown by loadings reflecting typhoid and tuberculosis control. Contingently, it justifies definition as that in the individual which generates community conscience and responsibility and which influences national statistics accordingly.

In addition to the foregoing, a further goal of the study will be to replicate major previously found factors for two reasons: (a) the analytical procedure to be pursued is that of factor-analytic simple structure which necessitates an adequate number of nonrelated variables to determine exactly the hyperplane of the factor under investigation *independently of the variables related to the hypotheses* and (b) the interrelationships of the present factor with other major dimensions is an important clarifying process. When one does find the factor replicated and clarified in the same total context as previously found, the chance of this being an artifact of the present research or its methods is decreased.

B. DEFINITION OF VARIABLES

The data collection is divided into three parts, each of which follows a separate criterion for finally selecting the variables to be used from among all those possible. The first category, primary-marker variables (variables all those possible. The first category, primary-marker variables (variables numbered 1 to 13 in this study), is that containing variables found in previous research to load a factor of morale, morality, and cultural integration and is used in this study to interlock it with previous concepts emerging for this factor. A second class of variables, numbered 14 to 29, consists of markers which have loaded *other* factors in previous studies and are included here not only as "hyperplane fodder" for good rotation of the morality factor, but also so that other factors which have appreciable influence on the primary marker variables will be recognized, if they occur.

The third class of variables covers the alternative directions in which we have speculated about the nature of the morale and morality factor per se. These variables, numbered 30 to 48 in the matrix list, are thus aimed to shed new light on the nature of the factor being investigated. The variables were chosen on the basis of their hypothesized importance and availability for all the countries. Many variables other than those in Table 1 were considered but were not included in the final analysis because of insufficient data available or because the ones included were thought to be more important. Unfortunately, in the end, a majority of our more "choice" variables were not reported by a large number of countries and therefore could not be included.³

The variables finally included, along with pertinent information about them, are as follows:

1. Syphilis death rate. The original raw-score distribution was posi-

³ The criterion for dropping a variable or country, a list of other variables, the correlations, the matrix of principal axis factors, the varimax solution, and other information are given in Gorsuch (13).

TABLE 1
RELIABILITY OF VARIABLES*

Variable No.	Title	Full-length reliability	Stability coefficient	Per cent "reliable"†‡§
1	Syphilis death rate	81	68	75
2	T. B. death rate	87	77	75
3	Birth rate	97	94	72
4	Alcohol death rate	95	90	75
5	Homicides and war death rate	84	73	70
6	Miles of railroad per inhabitant	99	99	***
7	Typhoid-fever death rate	96	92	75
8	Death rate (gross)	94	89	75
9	Marriage rate	95	91	84
10	People per dwelling	***	***	***
11	Population density	99	99	***
12	Students per teacher	80	67	***
13	Suicides	97	97	70
14	Area	99+	99+	99
15	Cities per inhabitant	***	***	***
16	Population	99+	99+	69
17	Clashes	37	23	***
18	Literature and science prizes	70	53	***
19	Government ministries	86	76	***
20	National income per inhabitant	97	95	***
21	Telephones per inhabitant	99+	99+	***
22	Science books	98	97	79
23	Cancer death rate	81	69	75
24	Buddhists	***	***	***
25	Heart death rate	91	84	75
26	Political parties	88	79	***
27	Defense expenditures	97	93	***
28	Mohammedans	***	***	***
29	Catholics	***	***	***
30	Climatic efficiency	***	***	***
31	Prostitution concern	***	***	***
32	Cost-of-living trend	***	***	***
33	Divorce rate	83	71	79
34	Forced labor	***	***	***
35	Hypertension death rate	74	58	70
36	Illegitimate births	99	99	82
37	International-aid index	***	***	***
38	News censorship	91	84	***
39	Peace prizes	84	72	***
40	Physicians per inhabitant	97	94	94
41	Population increase rate	***	***	***
42	Protestants	***	***	***
43	Religious books	89	81	77
44	Riots and rebellions	76	62	***
45	Students per inhabitant	93	87	***
46	UNICEF contributions per inhabitant	***	***	***
47	Latin American country	***	***	***
48	Extreme syphilis death rate	***	***	***
49	Random number	***	***	***
50	Random number	***	***	***
51	Random number	***	***	***

* All decimal points dropped.

** According to U.N. rating.

*** Not available.

tively skewed; therefore the data were normalized before correlation. It has loaded the factor under study in five separate investigations.

2. Tuberculosis death rate. Positively skewed, therefore normalized. Loaded the factor in three studies.

3. Birth rate (gross). Slightly bimodal; so normalized. Loaded the factor in three out of the four studies in which it appeared.

4. Cirrhosis-of-the-liver (alcoholism) death rate. Slightly positively skewed; so normalized. Loaded the factor in three out of five studies.

5. Homicide and war death rate. (Note inclusion of the latter; but if Richardson (22) is right, this contamination is not serious.) Positively skewed; so normalized. Loaded the factor in two out of three studies.

6. Miles of railroad per 100,000 inhabitants. Positively skewed; so normalized. Loaded in two out of three studies.

7. Typhoid-fever death rate. Positively skewed; so normalized. Loaded the factor in two studies.

8. Death rate (gross). Positively skewed; so normalized. Loaded the factor in one out of four studies.

9. Legal-marriage rate. Negatively skewed; so normalized. Loaded the factor in one study.

10. People per dwelling. Approximately normal; so normalized. Loaded the factor in one study.

11. Population density. Positively skewed; so normalized. Loaded the factor in one out of two studies.

12. Students per teacher. Primary school only. Approximately normal; so normalized. Loaded the factor in one study.

13. Suicide and self-inflicted-injury death rate. Positively skewed; so normalized. Loaded the factor in one out of three studies.

14. Gross area in 10,000 km.² Positively skewed; so normalized.

15. Cities over 20,000 per million inhabitants. Rectangular distribution.

16. Population (gross). In hundred thousands. Positively skewed; so normalized.

17. Clashes with other countries. Excludes wars; includes border conflicts involving armies. Continuing conflicts were scored once for each year. Positively skewed; so normalized.

18. Nobel Prizes in literature and science. Positively skewed with many zeros; so trichotomized.

19. Government Ministries. Approximately normal; so normalized.

20. National income per capita. In U.S. dollars. Positively skewed; so normalized.

21. Telephones per 10,000 inhabitants. Positively skewed; so normalized.

22. Science books published. Number of titles in pure and applied sciences. Positively skewed with many zeros; so normalized.

23. Cancer death rate. Approximately normal; so normalized.

24. Proportion of the population Buddhist. Many zeros; so dichotomized.

25. Heart-disease death rate. All types except hypertension with heart disease. Positively skewed; so normalized.

26. Number of political parties. Includes only those strong enough to elect candidates to national offices. Positively skewed; so normalized.

27. Proportion of national budget for defense. Colonies given one-half the score of the mother country. Positively skewed; so normalized.

28. Proportion of population Mohammedan. Many zeros; so dichotomized.

29. Proportion of population Catholics. Bimodal distribution.

30. Climatic efficiency. Huntington's rating of the human-work efficiency of the climates throughout the world, given the present civilization. Based on optimum temperature (extremes and averages), rainfall, and number of storm fronts as indicated by research. Average for the inhabitants according to 1950 population distribution. Rectangular.

31. Concern over traffic in persons and prostitution. Signers of the United Nations statement. Dichotomous.

32. Cost-of-living trend. Per cent increase of 1958 over 1953. Positively skewed; so normalized.

33. Divorce rate. Positively skewed with some zeros; so normalized.

34. Use of forced labor. For economic or political reasons. Dichotomous distribution.

35. Hypertension death rate. Approximately normal; so normalized.

36. Proportion of births illegitimate. Positively skewed; so normalized.

37. International-aid index. Includes all types of aid and loans in U.S. dollars. Base of 100. Less than 100 means aid received; more than 100, aid given. Approximately normal; so normalized.

38. News censorship. Rated according to censorship of outgoing news reports. Positively skewed with some zeros; so trichotomized.

39. Nobel peace prizes. Positively skewed with many zeros; so trichotomized.

40. Physicians per 100,000 Inhabitants. Bimodal.

41. Population increase rate. Per cent increase of 1958 over 1953. Approximately normal with slight bimodality; so normalized.

42. Proportion of population Protestant. Positively skewed with many zeros; so trichotomized.

43. Religious books published. Number of titles. Positively skewed; so normalized.

44. Riots and rebellions. Internal revolt or disorder of a serious and widespread nature. Continuing rebellions counted as two per year. Many zeros; so trichotomized.

45. Students per 1,000 Inhabitants. Approximately normal; so normalized.

46. Contributions to UNICEF per 1,000 Inhabitants. In U.S. dollars. Positively skewed with some zeros; so normalized.

47. Latin American Country. All Central and South American countries scored 1, all others scored 0.

48. Extreme syphilis death rate. All countries with a syphilis death rate greater than an absolute one *SD* from the mean of Variable 1 received a score of 1; all others were scored zero.

49. Random number. Positively skewed trichotomy.
50. Random number. Rectangular distribution.
51. Random number. Normal distribution.

The data in this study were collected from a period not overlapping that from which the previous definition of the morality factor was obtained (i.e., for the period of 1953 to 1958, with a few instances going back as far as 1947); but, of course, going back further for variables such as the total number of Nobel Peace Prizes. Data for each variable were, if possible, collected twice: once at the beginning of the time period and once at the end. Not only does this increase the stability of the results, but (what is more vital) it allows the calculation of stability and reliability coefficients (see Table 1). The United Nations has also evaluated the data published by its committees and has indicated the figures thought not to be reliable or to be inaccurate. In Table 1, the per cent of each variable which the United Nations indicated as being of lower reliability than the rest can also be found. The only variable of questioned reliability, in this sense, is number 17; but it was retained because of its importance as a marker variable.

C. THE SAMPLE OF COUNTRIES

Because of the possibility of capitalizing on chance in a correlation matrix, the common rule of thumb is to have at least three or four times as many observations (in this case, nations) as variables. To avoid difficulties from the number of variables approaching the number of cases in the special case when the total population of cases from which to draw is limited, one has two options: (a) including a small number of variables in the analysis, repeating the study several times with new extension variables each time, and comparing the results from several studies or (b) including all of them in one matrix and avoiding the problem of factor matching and crossinterpretation. The same number of variables is eventually involved in either case, except the latter method is quicker and makes use of more information. It is obvious that the more information one has to base conclusions on, the more likely the conclusions are to be correct, even with error entering into each bit of information, *assuming the same number of conclusions are drawn in each case*. Also the present study is a special case since the sample, although small ($N = 52$), contains virtually all of the population; thus eliminating the ordinary generalization problems (except, of course, in terms of generalizing over time). However, the reader must be cautioned not to draw wider conclusions than the sample warrants.

Another caution and limitation is placed on the present research by the

nature of the variables. The data were usually reported by individual nations and many sources of error can enter, including chance miscalculation, defects in the data from improper sampling, differing definitions of the variables, and of course from motivation distortion through political aim or pride.

The nations on which sufficient data were available are as follows: Argentina, Australia, Austria, Belgium, Brazil, British Guiana, Burma, Canada, Ceylon, Chile, China (Taiwan), Columbia, Costa Rica, Czechoslovakia, Denmark, Dominican Republic, Ecuador, Egypt, Finland, France, Germany (West), Greece, Guatemala, Hungary, Iceland, Ireland (Irish Free State), Israel, Italy, Japan, Luxembourg, Maltese Islands, Mauritius, Mexico, Netherlands, New Zealand, Nicaragua, Nigeria, Norway, Panama, Portugal, Puerto Rico, Rhodesia and Nyasaland, Spain, Sweden, Switzerland, Trinidad and Tobago, Union of South Africa, Union of Soviet Socialist Republics, United Kingdom, United States of America, Uruguay, and Yugoslavia (a total of 52).

D. THE ANALYSIS

The Pearson product-moment formula was used to calculate the correlations among the 51 variables. Its advantages for such work are that it utilizes the maximum amount of information and may be computed on any distribution that is fairly symmetrical and unimodal so long as it is not badly skewed. Applying the Pearson formula to the dichotomies and trichotomies gives comparable coefficients to those obtained from the other variables (14). The major disadvantage of the reduction of some variables to dichotomies and trichotomies is the lost information and therefore the greater inaccuracy of the coefficients; however including the variables in their original forms was not an acceptable option because of the extreme skewness of the variables (the usual transformations to eliminate skewness could not be applied because of numerous zeros that would be unaffected); therefore the present solution provides a practical approach to these variables. After examination of the distributions, the authors decided to transform the majority of the variables to a normalized condition which most of them closely resembled. Scatter plots were checked in regard to the assumption of linearity.⁴ The variables were

⁴ From the examination of scatter plots, two variables were seen to have some curvilinearity with other variables, although it was slight. The first of these, syphilis death rate (Variable 1) had several relationships showing just-noticeable curved regression; therefore it was entered again into the correlation matrix but scored this time in terms of the extremeness of the death rate in either a large or small direction from the mean and is Variable 48. One other variable, clashes (Variable 17) also appeared to have several curved regressions from the influence of a dozen relatively homogeneous countries. Therefore, Variable 47 (Latin American countries) was in-

then correlated with the aid of the Illiac digital computer; the resulting 1,275 correlations are given elsewhere (13).

The Guttman criterion of temporarily putting unities in place of communalities and stopping extraction when a latent root first drops below unity was used to decide the number of factors. Communalities were then estimated iteratively for the 14 factors indicated. The unrotated matrix is given elsewhere (13) along with the residuals after 14 factors were extracted by the principle-axes method. Cattell's definition (8) of Thurstone's simple-structure criterion—the maximized total per cent of zero loadings—was used to direct visual rotations to a unique resolution. Analytical approximations were first tried (by both orthogonal and oblique programs), but were found to give poorer simple structure than the visually rotated solution. After seven hand rotations, reaching a plateau with a ± 10 hyperplane count of 58 per cent, it was felt that the simple-structure criterion had been met. The unrotated matrix and the reference vector structure from which the hyperplane count was made are given by Gorsuch (13). The *primary-factor pattern*, with the factors arranged in the approximate decreasing order of variance accounted for by each factor, is given in Table 2, and Table 3 gives the correlations among the factors. The loadings of the random variables gives an approximate level of "significant loadings." All calculations were to at least five digits and rounded only when placed in the tables.

For the interpretative discussion in the following section, it has seemed best to set out the other factors as well as our main concern—the morality factor. Confidence in a factor rests on the framework of known factors about it also being replicated and, in this case, the obvious match of most of these to previous factors justifies higher confidence in our main result. Evaluations of goodness of matching were made by the Cattell-and-Baggaley salient-variable similarity index, s (7), the values of which are discussed under each factor.

E. VERIFICATION OF CULTURE PATTERN DIMENSIONS

The pattern of the main morale and morality factor (Factor 3 in the present study) clearly matched with the original one so identified (3). Its major-loading variables, given in rank order, are as follows (with adjectives describing the high end of the factor and the number of the variable given in parentheses): low syphilis death rate (1), low tuberculosis death rate (2),

cluded in the analysis. In both cases, the original variable was kept in the analysis also. Since Variable 48 has not loaded any factor to any important extent, the curvilinear relationships of Variable 1 are negligible. Variable 47 loaded only one factor, Size; therefore it appears that the curvilinear relationships for Variable 17 are limited.

TABLE 2
PRIMARY-FACTOR PATTERN*

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	-02	-19	-86	05	14	-13	01	-58	-03	-04	01	-50	-06	25
2	-62	17	-79	-05	03	-07	-12	-48	09	09	-46	-26	-05	15
3	02	48	-00	50	15	05	-19	00	-03	-48	-06	-32	-09	19
4	-25	-02	-51	00	-04	08	-35	05	-01	-09	04	30	09	44
5	-43	67	-47	10	21	07	-23	03	-23	-41	02	18	00	-15
6	94	06	-18	-87	41	08	12	50	-08	-07	11	13	-11	-08
7	-62	23	-44	05	18	-12	-05	05	-29	-24	05	-16	15	08
8	23	54	-11	-40	24	07	11	14	01	-33	-45	-18	11	-32
9	-06	-50	-21	19	-03	-42	-37	-60	19	19	11	11	01	07
10	-77	40	-23	48	18	13	07	15	04	06	-01	08	10	34
11	-64	-81	-05	24	-36	-08	10	-95	43	-09	-05	00	-04	-03
12	-47	-02	-08	11	-01	-05	-37	-07	01	00	-35	03	-47	-05
13	06	-04	-08	-64	-06	17	-12	00	01	83	07	16	39	08
14	57	-04	10	-19	86	17	-53	95	-14	-15	-07	-05	-05	19
15	36	-08	07	13	-46	02	-29	09	+1	-07	36	-42	-03	03
16	-09	-71	-01	00	84	10	-56	12	27	02	-09	-08	05	10
17	-47	23	-07	13	18	-38	-10	02	71	11	20	15	05	07
18	11	-44	11	-19	24	00	-20	14	32	44	12	-06	05	-12
19	00	-85	08	-10	63	12	03	09	10	-04	00	-10	-06	02
20	94	-21	10	-48	-17	08	-04	07	08	07	48	-02	-11	02
21	89	-23	04	-48	00	13	-05	31	05	20	31	13	-09	00
22	-13	-58	-03	-23	34	-28	-35	-24	28	63	11	-24	11	-02
23	34	-19	-00	-68	-10	07	12	09	11	50	05	35	05	-22
24	-49	-56	12	05	-15	15	-67	-09	29	-04	-01	-06	00	88
25	64	-60	-08	-77	-03	05	08	17	00	-04	01	18	-23	00
26	04	-52	-03	15	27	111	23	-34	-09	26	11	51	43	03
27	09	08	-11	61	00	02	-69	-05	08	31	07	05	-16	-24
28	00	-56	51	08	10	-03	-49	-10	-44	-17	12	13	08	10
29	-23	50	-74	-06	00	-13	61	-09	10	-33	06	-14	-04	-19
30	25	-29	-08	-58	-08	-33	10	09	37	11	-06	00	00	-07
31	-36	-15	-23	-47	07	-18	-04	-02	-01	07	-12	02	47	61
32	-13	02	-22	31	10	63	-01	41	-06	-11	-33	85	-10	-08

* See text, pp. 80 to 83, for definitions of variables.

TABLE 2 (continued)

Variable	Factor													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
33	41	15	44	-06	06	-22	-23	-50	-05	78	-01	-02	26	-37
34	-41	-40	05	-45	-01	-105	-01	12	04	02	06	-38	-18	-10
35	49	-50	08	-23	-03	07	03	03	-09	-20	19	43	-05	-10
36	02	98	-55	-35	10	10	03	15	-46	80	09	-53	-09	-07
37	77	-79	-03	-34	13	08	-31	14	-01	-30	-09	-34	-34	10
38	-75	02	-18	02	10	-103	-11	-09	01	15	-07	-16	-11	-34
39	64	-08	11	-02	38	42	-11	-15	18	44	21	-08	19	-51
40	02	-75	-10	-15	-01	-18	05	-03	08	04	50	40	18	-05
41	06	100	08	52	-03	13	-05	-09	00	17	33	03	23	-13
42	94	27	56	-60	13	23	08	31	-05	50	-07	-10	09	-03
43	-06	-79	-07	-11	52	13	-26	-34	20	52	04	02	10	-13
44	-58	46	-11	03	28	00	-01	45	00	19	-27	18	-23	-37
45	13	-87	-04	-16	-11	28	-10	-45	24	06	33	16	-37	39
46	91	-06	-15	06	58	46	-04	-36	10	06	02	00	-11	-19
47	02	17	-19	06	56	00	-10	28	-05	-06	42	-40	09	17
48	09	10	07	15	08	32	-05	-05	-34	-39	-05	-22	-03	00
49	-26	-19	-27	47	-22	-10	-03	02	-04	-44	-19	13	15	-03
50	-28	29	32	21	04	13	04	50	58	06	-09	-14	08	07
51	-15	-12	44	-03	01	-25	14	00	05	46	-04	02	-43	-19

TABLE 3
PRIMARY-FACTOR CORRELATIONS*

Factor	1	2	3	4	5	6	Factor 7	8	9	10	11	12	13	14
1														
2	—60													
3	—48	52												
4	60	—49	—47											
5	—22	50	17	—17										
6	—54	21	31	—44	—28									
7	44	—39	—24	50	—21	—16								
8	36	—67	—54	51	—36	—10	39							
9	24	—38	—35	01	—20	07	—02	32						
10	69	—67	—54	64	—25	—41	48	59	30					
11	—35	35	13	—27	01	27	35	—20	—03	—22				
12	34	20	—01	16	45	—59	07	—30	—17	12	—10			
13	22	—42	—17	05	—19	—15	—07	26	20	09	—37	—12		
14	20	30	08	27	39	—48	22	—29	—39	11	09	60	—29	

* See text, pp. 85 to 95, for definitions of factors.

small percentage of population Catholics (29), large percentage of population Protestants (42), small proportion of births illegitimate (36), low alcohol death rate (4), large percentage of population Mohammedans (28), low homicide and war death rate (5), low typhoid-fever death rate (7), and high divorce rate (33). In the present study, it is being called *Morality-Responsibility vs. Poor Cultural Integration*.

When this factor is statistically matched with Cattell's⁵ Factor 12 on 20 comparable variables, the s index gives significant match ($p < .05$). Beyond this, the "subjective" judgment of similarity is also strong. However, the factor has been simplified through several variables (6 and 10) not loading the present factor. These had led to an earlier hypothesis of the factor being related to slum conditions, but the connection is slighter now. (Incidentally, no importance should be given to the fact that the factor here ranks third now in terms of variance accounted instead of 12th as in Cattell since we deliberately limited the study here primarily to the variables in which this factor plays a major role.)

There are several hypotheses that, from a superficial examination of the variables, could easily mislead the unwary. As Cattell (3, p. 466) earlier noted, the factor must not be confused with "health influences" for two reasons: (a) in this study, many variables such as death rate and physician rate load a distinct factor, whereas it would be imperative for them to load this one if the factor was concerned specifically with health measures, and (b) the death rates related to this factor are specifically only those that medical science can greatly alleviate provided the individual or society is sufficiently motivated to act. However, under conditions of low morale, the persons responsible and their community supporters do not act in accord with these responsibilities and the high death rates follow—not to mention a decrease in synergy.

Of the three death rates, two (syphilis and alcoholism) can be easily prevented through appropriate abstinence or through proper treatment available in many of the countries studied; therefore it appears that some lack the foresight, self-responsibility, and internalized controls to act in a preventive manner. This, of course, may be the result of the culture's general approach to areas needing such responsibility.

The tuberculosis death rate is twice affected by a person's acceptance of responsibility. First, the incident rate is influenced by personal hygiene, particularly as it affects general resistance and the carefulness a person

⁵ All matches with and references to previous factors in this section are to the original factors as given in Cattell (3), except when otherwise noted.

exhibits when in possible contact with the *tubercle bacilli*. Secondly, and more important for this factor, tuberculosis is a chronic relapsing disease, such that if a person does not strictly follow medical advice, a relapse and possibly death will result. (Note that the other factors have essentially removed all the variance in these variables attributable to economic and medical development; therefore, these two interpretations of these particular variables cannot be used here.) Thus, in this death rate there also appears to be a personality element involving the related concepts of internalized controls, accepting responsibility for oneself, and utilizing self-control to follow the best course. Similar interpretations can be given to many of the other variables.

Other variables one would expect to load the factor under the interpretation developed here usually have the correct sign and are out of the hyperplane but not sufficiently so to be salient with a small sample.

A question that immediately arises is whether this is a *population* characteristic of "general morality level," or a *syntality* characteristic of "morale," where the members of the group uphold the standards because of the group atmosphere. These have been sometimes confused in many of the writings on group dynamics, where measures of individual verbal attitudes have been taken as identical with a group syntality measure. Granted that most theorists consider morality and morale distinct concepts, let us state our present hypothesis on this distinction.

The analyses previously reported and the present one have not permitted, as confessed, anything like an adequate degree of hypothesis-directed control of the choice of variables in relation to morale hypotheses. Consequently, our theory of morale rests primarily on the small-group experiments. In writings on morale generally it has meant anything from euphoria to normality and from group effectiveness to conformity. As it derives from the Cattell and Stice, Gibb, and Lawson studies (9, 12, 19), it means primarily the magnitude of total synergy (determining especially cohesion, the group resistance to its disruption) as contributed to largely by the distinct leadership (personal or ideational values) and social (gregariousness or "buddy" satisfaction) primary factors. At present the nature of the *remaining* three or so morale primaries is not interpreted, but the "fortitude" factor is shown by its correlations to be tied to qualities of personality (16 PF measures) in the population.

Since small groups taken for experiments are always taken out of large cultural groups and borrow their morality from the cultures of the latter (in the superego habits of the constituent members), the resulting smallness of variance, complicated by the poor choice of variables used, may account for the

experimental failure to find in small groups the morality factor found in large groups. Alternatively, however, one may argue that the distinctive *leadership* synergy factor is the small-group equivalent of large-group morality, provided we include leadership of an idea or value in the former.

By either approach morality, as "population conformity to ideals," is only one of the full range of factors that contributes to the group behaviors commonly included under the term "morale." At the same time, acceptance by the population of values of self-sacrifice, the Ten Commandments, and other values of altruistic morality probably play a definite role in the success of the group and its resistance to disruption: i.e., to the totality of morale. This theory that despite the variations in religious form of universalist ethics those ethics have essentially developed as intergroup survival adaptations within large groups has been worked out more fully by Cattell (2). The factor found above in national groups clearly appears to correspond to this *morality* contribution to morale—not to *all* influences favoring group survival. Even so, it deserves investigation without prejudice that its empirical form will be restricted to manifestations fitting the dictionary definition of morality. In particular, it may include further features, such as higher group integration, cultural consistency, and the social conscience of a group regarding its members.

The definition of the morality dimension in factor form makes possible a weighted composite to measure this aspect of cultures. From an averaging of the several studies mentioned, with particular emphasis on the present one, we arrive at the following for use with standard scores to determine *low* morality: syphilis death rate, .6; alcoholic death rate, .4; homicide death rate, .4; proportion of births (illegitimate), .2; people per dwelling, .2; and tuberculosis death rate .1.

The foregoing weights were subjectively determined on the basis of consistency of loading in the studies in which the variable appeared, magnitude of loading, and calculation of multiple-regression coefficients on one sample (which showed, for example, that the tuberculosis death rate's contribution is predominately accounted for in the same variance as the syphilis death rate; thus necessitating a lower weight than otherwise would be accorded to it). Variables loading at a borderline level in the present study have been included only if they were salient in other studies. Naturally, the weighting system is tentative and must be improved through further study.

When countries were scored on the four most-salient variables of this factor (in their standard-score form), Brazil, Burma, Dominican Republic, Mexico, and Portugal were found among those low; while Denmark, Iceland,

Israel, and Sweden were prominent at the high end of the factor. From a world-reputation viewpoint, the former countries fit the suggested syndrome of low morale, immorality, and denial of responsibility more than the latter.

From the primary factor correlations of Table 3, one sees this factor forming a cluster with Factor 2 positively and with Factors 8 and 10 negatively, confirming some of the previously found relationships; the lack of cultural integration is a possible basis of the cluster.

Factor 1, Vigorous Development *vs.* Lack of Exploitation of Resources, confirms Cattell's factor of "Vigorous, Self-Willed Order *vs.* Unadapted Perseveration," with a match at the .03 probability level by the *s* index. The variables loading it are many miles of railroad per capita (6), high national income per capita (20), large percentage of population Protestants (42), large UNICEF contributions per capita (46), many telephones per capita (21), few people per dwelling (10), high international-aid index (37), and little news censorship (38). While poorly established in Cattell's study, it is clearly here one of the general control of environment based on a well-developed technology and, from the low news-censorship loading, a society not afraid of free exchange of ideas. The indications are that it is a result of something akin to a high self-assertive erg on a national level and is intrinsically inherent in high group effectiveness.

Whereas Factor 1 related to technological development, Factor 2 (Underdeveloped "Backwardness" *vs.* Complex Social Systems) appears to relate to the development of "culture" in the narrower sense in the people themselves. Its major variables are large population increase rate (41), large proportion of births illegitimate (36), few students per capita (45), few government ministries (19), low population density (11), low international-aid index (37), few religious books published (43), and low physician rate (40). The high end of the continuum would be typified by the ignorant, backward people who have little contact with the government (e.g., illegitimate births is a function of morality *and* the extent to which the government legalizes and records marriages throughout the country). Cattell's Factor 10 (Fastidiousness *vs.* Forcefulness) is a *tentative* match for this factor.

Except for a suggested slight change in the previous factor's title, "Patriarchal Solidarity *vs.* Ferment of Release" (which matches in actual pattern with the present factor at the .05 probability level), Factor 4, Patriarchal Solidarity *vs.* Stimulating Atmosphere, is basically unchanged from Cattell's earlier factor, with its variables being few miles of railroad per capita (6), low heart-disease death rate (25), low cancer death rate (23), low suicide death rate (13), high percentage of government budget for

defense (27), low percentage of population Protestants (42), and low climatic efficiency (30). Here Cattell's earlier interpretation is not only confirmed but also expanded. The death rates show a lack of stress, and the climatic-efficiency variable indicates that the countries high on this factor are those in which the climate discourages high standards of living. The factor is in keeping with Huntington's hypotheses (17) concerning this variable.

Factor 5, Size, is defined by the following variables: large area (14), large population (16), many government ministries (19), large UNICEF contribution per capita (46), many Latin American countries (47), and many religious books published (43). Both the match to Cattell's first factor of Size (at the .01 level), and the "size" interpretation of Factor 5 cannot be doubted. Two of the variables show interesting effects: (a) UNICEF contributions to feed starving children show what one might call the "Big Brother" effect of the modern political climate and (b) religious books are related to population size but science books seem, at present, not to be; thus showing a discrepancy between interest in science and religion on a world-wide view.

Factor 6, Democracy *vs.* Totalitarianism, is loaded by these variables: many political parties (26), little use of forced labor (34), little news censorship (38), and upward cost-of-living trend (32). Although there is not a match in any earlier study for this factor, due to a lack of markers, one can easily accept its reality in terms of new but obvious interpretations. It is at this point that we would have liked to have been able to interview natives in relation to morale.

Another clear factor, matching with Cattell's factor of Buddhist-Mongolism—though it happens to come reversed—at the .01 level, is Factor 7: Occidental *vs.* Buddhist-Mongolism, loaded by small percentage of government budget for defense (27), small percentage of population Buddhist (24), large percentage of population Catholic (29), small population (16), small area (14), and small percentage of population Mohammedan (28). The religions loading this factor tend not only to be mutually exclusive but also to exert their major influences in different parts of the world; moreover it places the small Western country, which appears to be relatively homogeneous, in opposition to the larger but also homogeneous countries of the East. Therefore, we reject an exclusively religious interpretation in favor of a broader cultural interpretation that includes the religious elements.

A probability at the $p = .05$ level is found for the match between Factor 8, Narrowness *vs.* Enlightened Affluence, and Cattell's factor of "Enlightened

Affluence *vs.* Narrow Poverty," although of course reversed in direction. The following variables are salient: low population density (11), large area (14), and low marriage rate (9). What is surprising is its smallness in the present study since it was one of the better determined and larger factors in the original work. Two hypotheses, which present evidence cannot decide between, can be offered for this change. The first and most probable is that our selection of variables eliminated many of those related to this factor. The second hypothesis is that since the time of the data-gathering periods of previous studies, which mostly used data collected before World War II, the variance of this factor has decreased as more countries gained the minimum technological development to begin providing schools, etc. The area seems ripe for more extensive study.

Factor 9, Cultural Pressure *vs.* Direct Expression of Drives, is loaded by many clashes (17), random variable (50), small proportion of births illegitimate (36), and small percentage of population Mohammedan (28). It has the only major loading of a random variable (which warns one not to put too much faith in any single variable's loading) but is still clearly the same factor Cattell found (matches at the .01 level). It seems to differ from Factor 4, Solidarity *vs.* Stimulating Atmosphere, in being cultural; whereas the cultural ferment in Factor 4 appears to be more dynamic.

Factor 10, Anomie *vs.* Channeled Emotionality, has the following variables salient: high suicide death rate (13), large proportion of births illegitimate (36), and high divorce rate (33). It bears a striking resemblance to some of Durkheim's (11) work on suicide rates in relation to divorce, etc. His general theory states that suicides result from "anomie" since the society can no longer serve its normal function of channeling the person's emotions into suitable outlets and social solidarities. The extra tension produces a situation which often becomes intolerable to the person and results in suicide. Its *s* index match is at the .05 level with Cattell's factor which places the more rational and scientific aspects of life in opposition to the more emotional ones.

Factor 11, Medical Development, is another factor without a match in Cattell's analysis, apparently because the necessary data were not available there; it is clearly the result of health through proper treatment, which the countries with sufficient resources are able to afford. Its salient variables are high physician rate (40), high national income per capita (20), low tuberculosis death rate (2), and low death rate (8).

The last three factors are, in this study, too small to interpret with any reasonable degree of confidence. Factor 12 matches (at the .03 level) with

Cattell's Emancipated Urban Rationalism *vs.* Unsophisticated Stability and Factor 13 with Bourgeois Philistinism *vs.* Reckless Bohemianism (.03 level).

F. SUMMARY

This research aimed to clarify a cultural dimension of morale and morality found in rough form in several earlier factor analyses of national statistics. Data on 48 variables for 52 countries in the early 1950s were correlated and factored (with three random variables) yielding 14 factors rotated to simple structure.

The majority of these factors match statistically in pattern and in associated meaning the factors found by Cattell on data of the 1930s and 1940s; though two comparatively clear new factors are added, tentatively named Democracy *vs.* Totalitarianism and Medical Development. Although our new variables enriched the other factors, their main contribution was to the morality factor. Since this factor repeats its pattern at a significant level ($p < .05$) and stands in a framework of nine other factors just as the former factor did, its identity as a distinct factor is supported.

The meaning of the morality factor is conceptually distinct as a component of morale, which is probably a second-order factor; the present evidence reduces its association with slum conditions or health conditions *per se*, and its absence of loading on school education suggests it is an aspect of cultural values passed on more through the family, the peer group, or religious organizations. Central in it is a sense of community responsibility and superindividual dedication, probably not transmitted through laws and outer controls but through inner values.

REFERENCES

1. BENEDICT, R. *Patterns of Culture*. Boston: Houghton, 1934.
2. CATTELL, R. B. *Psychology and the Religious Quest*. New York: Nelson, 1938.
3. ———. The dimensions of culture patterns by factorization of national character. *J. Abn. & Soc. Psychol.*, 1949, **44**, 443-469.
4. ———. A quantitative analysis of the changes in culture patterns of Great Britain, 1837-1937, by P-technique. *Acta Psychologica*, 1953, **9**, 99-121.
5. ———. Group theory, personality and role: A model for experimental researches. In F. A. Geldard (Ed.), *Defence Psychology*. New York: Pergamon, 1961.
6. CATTELL, R. B., & ADELSON, M. The dimensions of social change in the U.S.A. as determined by P-technique. *Soc. Forces*, 1951, **30**, 190-201.
7. CATTELL, R. B., & BAGGALEY, A. R. The salient variable similarity index for factor matching. *Brit. J. Stat. Psychol.*, 1960, **13**, 33-46.
8. CATTELL, R. B., & GORSUCH, R. L. The uniqueness and significance of simple structure demonstrated by contrasting organic "natural structure" and "random structure" data. *Psychometrika*, 1963, **28**, 55-67.

9. CATTELL, R. B., & STICE, G. F. The dimensions of groups and their relations to the behavior of members (2nd ed.). Champaign, Ill.: Laboratory of Personality Assessment and Group Behavior, 1960.
10. CATTELL, R. B., BREUL, H., & HARTMAN, H. P. An attempt at a more refined definition of the cultural dimensions of syntality in modern nations. *Amer. Sociol. Rev.*, 1952, 17, 404-421.
11. DURKHEIM, E. Suicide. Glencoe, Ill.: Free Press, 1951 (Original, 1897).
12. GIBB, C. A. Changes in the culture pattern of Australia, 1906-1946, as determined by P-technique. *J. Soc. Psychol.*, 1956, 43, 225-238.
13. GORSUCH, R. L. National morale, morality and cultural integration. Unpublished Master's thesis, University of Illinois, Urbana, 1962.
14. GUILFORD, J. P. Fundamental Statistics in Psychology. New York: McGraw-Hill, 1956.
15. HARTSHORNE, H., & MAY, M. Studies in the Nature of Character. New York: Macmillan, 1928.
16. HOFSTAETTER, P. R. A factorial study of cultural patterns in the U.S. *J. of Psychol.*, 1951, 32, 99-113.
17. HUNTINGTON, E. Mainsprings of Civilization. New York: Wiley, 1945.
18. KEYSERLING, H. Das Spektrum Europe. Stuttgart: Deutsche Verlags-Austalt, 1928.
19. LAWSON, E., & CATTELL, R. B. Sex differences in small group performances. *J. Soc. Psychol.*, 1962, 58, 141-145.
20. MEAD, M. Cultural Patterns and Technological Change. New York: New American Library, 1955.
21. ———. Research in contemporary cultures. In H. Guetzkow (Ed.), *Groups, Leadership and Men*. Pittsburgh: Carnegie Press, 1951.
22. RICHARDSON, L. F. Distribution of wars in time. *J. Royal Stat. Society*, 1946, 107, 242-250.
23. THORNDIKE, E. L. Your City. New York: Harcourt, 1939.
24. TOYNBEE, A. J. A Study of History. New York: Oxford Univ. Press, 1947.

Department of Psychology
University of Illinois
Urbana, Illinois

Department of Psychology
Vanderbilt University
Nashville 5, Tennessee

SOCIAL-DESIRABILITY RATINGS OF PERSONALITY ITEMS BY THREE SUBCULTURAL GROUPS*¹

Department of Psychology, University of British Columbia

CAROL JEAN DIERS

A. INTRODUCTION

In constructing his Personal Preference Schedule, Edwards (2) attempted to control the tendency for subjects to respond to personality questionnaires in a socially desirable manner. Items in the EPPS were scaled for degree of social desirability and two items, each representing a different personality variable, were paired in terms of similar social-desirability scale values. In taking the schedule, the testee is forced to make a choice between paired items.

Judgments of how desirable or undesirable various groups consider certain attitudes or behavior in others have been obtained to determine the generality of these ratings (3, 6). Intercorrelations of social-desirability ratings of EPPS items are consistently positive and significant. This has led to the suggestion that there is a social-desirability stereotype that persists across many subcultural groups. If such were the case then tests, such as the EPPS, may be constructed to control the social-desirability variable in personality testing.

In the case of the EPPS, an additional examination of the intraclass correlation between matched pairs is required before one can conclude that the social-desirability variable is controlled for a given group. Edwards (2) reports an intraclass correlation of .85 for social-desirability scale values used in constructing the EPPS. Klett (5) reports the correlation between matched pairs for high-school students as .69 and Fujita (4) found an intraclass correlation of .76 for the Japanese-American group.

Klett and Yaukey (6) summarize social-desirability ratings of EPPS items made by six different groups. They indicate that even though there are high positive correlations on the judgments by various groups there may be

* Received in the Editorial Office, Provincetown, Massachusetts, on April 6, 1964. Copyright, 1965, by The Journal Press.

¹ Based on a portion of the author's Master's thesis. The research was supported in part by a grant from the President's Fund, University of British Columbia, under the direction of Dr. Douglas T. Kenney. The author is indebted to Dr. Kenney for his assistance and guidance throughout the course of the study.

important systematic differences when similar items are grouped according to the 14 "needs" the EPPS purportedly measures. Klett and Yaukey use these differences in need rating to infer similar differences in the groups making the rating.

This study reports analyses of judgments made by three new groups. The first hypothesis was that intercorrelations of judgments made by the three new groups with those made by the six groups reported by Klett and Yaukey would be positively significant. The second hypothesis was that intraclass correlation of item pairs for the three new groups would be positively significant to a degree warranting the assumption that the social-desirability variable would be controlled for these groups. The third hypothesis was that these three groups and Edwards' original University of Washington sample would not differ in mean rating of the 14 EPPS needs.

B. METHOD

Social-desirability judgments of the 135 statements in the EPPS were obtained from (a) University of British Columbia students, male and female; (b) Hungarian immigrant university students, male²; and (c) Canadian juvenile delinquents, female.³

The Canadian university sample consisted of 123 males and 103 females enrolled in two sections of introductory psychology at the University of British Columbia. The age range for the 226 subjects was 17 to 36, with a mean age of 21.2 and a standard deviation of 3.8 for the males, and a mean age of 18.8 and a standard deviation of 2.1 for the females.

The Hungarians were 70 male students who were part of the University of British Columbia, Faculty of Forestry, Sopron Division. Shortly after the Soviet-occupation authorities employed armed force on November 4, 1956, to suppress the October 23, 1956, Hungarian revolution, the faculty and students of the School of Forestry, located in Sopron, left Hungary as a group. After a brief stay in Austria, this group immigrated to Canada and resumed their studies at the University of British Columbia. The major portion of their university classes were taught by the original Sopron faculty and except for special classes in English the classes were taught in Hungarian. Retention of the internal group structure has no doubt retarded the assimilation of individual Hungarian students into Canadian society. The group had

² Dean R. G. Roller of the Sopron School of Forestry was very helpful in permitting testing of the Hungarian students.

³ The author is indebted to Miss D. Johnstone, Treatment Director, Girls Industrial Home for permission to test the juvenile delinquents.

been in Canada 14 months at the time the social-desirability scale values were obtained. It is worth noting that the official policy at the Sopron School of Forestry was to admit 70 per cent of the students from the peasant and worker classes and 30 per cent from the intellectual and white-collar classes. The immigrant group consisted of a slightly higher percentage of individuals from the white-collar classes than this expressed ideal (7). The age range of the 70 Hungarian students was from 20 to 32 years, with a mean of 22.7 and a standard deviation of 2.6.

The delinquent group consisted of 40 institutionalized females in the Girls Industrial Home, Vancouver, B. C. The most frequent charge against this sample was "incorrigibility," a legal term covering most types of misdemeanors, other than that of sex offenses, which was the second most-frequent charge. According to the Treatment Director, most of the girls were from the lower socioeconomic strata and about one-half were wards of the province or some private children's agency. The age range of the 40 delinquents was 13 to 18 years, with a mean of 15.7 and a standard deviation of 1.2.

All subjects rated the items on a nine-point scale of social desirability according to instructions outlined by Edwards (3). The Canadian university students gave their ratings during a regular class hour, as did the Hungarians. Instructions and statements were administered in Magyar to the Hungarians.⁴ The juvenile-delinquent ratings were obtained in an assembly called for that purpose.

C. ANALYSES AND RESULTS

Social-desirability scale values were computed by the method of successive intervals. Male and female ratings given by the Canadian university sample were first computed separately. The correlation between the two sets of ratings was .95. Successive-interval ratings were then obtained for the combined sample of 226.

The ratings obtained from the three groups were correlated with all other ratings reported by Klett and Yaukey,⁵ as shown in Table 1.

All correlations are significant at the .01 level. (Nine items from need abasement are not included because data were not available from all samples.) The result in the main substantiate the findings of previous investigations; that is, persisting across several subcultural groups is a common agreement regarding socially desirable personality characteristics. Excluding Hungarian

⁴ Thanks are due Dr. J. Huberman for translation into Magyar and administration of the Hungarian rating scale.

⁵ The ratings were provided by C. J. Klett, personal communication, 1956.

TABLE 1
INTERCORRELATION OF SOCIAL-DESIRABILITY SCALE VALUES FOR CANADIAN,
HUNGARIAN, AND DELINQUENT SAMPLES WITH SIX OTHER GROUPS

Group	Canadian University	Hungarian	Delinquent
American University	.95	.71	.58
Hospital	.81	.73	.50
High School	.90	.65	.62
Beirut	.82	.78	.54
Norway	.80	.71	.62
Nisei	.93	.69	.51
Delinquent	.54	.47	
Hungarian	.72		.47

and delinquent ratings, the Canadian university correlations range from .80 to .95. This is comparable to previous reports. However, the other two groups in this study (while providing significantly positive correlations) are more deviant than any previously reported. The Hungarian correlations range from .65 to .78. All the correlations with the delinquent sample are lower than those thus far obtained, the range being .47 to .62.

Matched-pair correlations indicate the degree to which the forced-choice items would be equated for social-desirability scale value if the EPPS were used on these populations. An intraclass correlation of .73 was obtained from the Canadian university sample. While this correlation is significantly different (at $p < .05$) from the original matched-pair correlation of .85 obtained by Edwards (2), nevertheless a good degree of covariation is evident. The intraclass correlations of .38 obtained from the Hungarian group and that of .35 for the delinquent group do not seem adequate to warrant the assumption that the EPPS controls the set to respond in a socially desirable manner for these populations.

Item ratings by the three new groups plus Edwards' original sample of 152 American university students were grouped according to the 14 needs, and analyses of variance were performed to compare mean need rating. In order to eliminate possible rating-scale bias, each of the four sets of scale values was transformed into normal deviates with a mean of 50 and a standard deviation of 10. Means and variances were thus equated on the total set of items, but individual items within each group were left free to vary. The 135 items were then grouped into the 14 needs, and mean ratings on each of the needs were computed for each group. The analyses indicated whether differences in item ratings were confined to a specific need or could be accounted for in terms of random variation.

Eleven of the 14 needs showed significant differences. Deference, succor-

ance, and nurturance produced nonsignificant F ratios. Mean ratings on exhibitionism, affiliation, and dominance were significantly different (at the .05 level). The remaining needs—achievement, order, autonomy, intraception, change, endurance, heterosexuality, and aggression—varied significantly in mean rating (at the .01 level).

Duncan's multiple-range test (1) was employed in the 11 needs that resulted in significant F ratios to determine the specific group or groups contributing to the differences. The group means on each of the "needs" and significant differences in mean rating are presented in Table 2.

There are no indications that American and Canadian students differ significantly on any of the needs rated. Further, Hungarian, Canadian, and American students regard the needs of achievement, exhibitionism, intraception, change, endurance, and heterosexuality in a similar manner. Hungarians regard the need affiliation as less desirable than do the American and Canadian college students and Hungarians rate the need dominance as less desirable than do Canadian students. The need order is rated higher by the Hungarian group than by any others.

Compared to the three other groups, Canadian delinquents rate the need order, intraception, and endurance significantly lower in desirability. Compared to American and Canadian college students, delinquents undervalue the need achievement. The needs exhibitionism, autonomy, change, heterosexuality, and aggression are rated as more desirable by the delinquents than by the Canadian group. Of these five needs, the delinquents view the needs autonomy, change, and aggression as more desirable than do the American college students; and the needs exhibitionism, change, and heterosexuality as more desirable than do the Hungarians.

D. DISCUSSION

It should be noted that evaluation of a trait as being socially desirable does not necessarily indicate a corresponding high need; for example, a group may rate aggression as being socially desirable, but that same group may or may not have high needs for aggression. The data presented merely indicate which traits certain groups rate as being more or less socially desirable than other groups.

Canadian and American college students, who show no differences in their need ratings, will be taken as referent points for discussing the Hungarian and delinquent ratings.

Hungarians value the need order more highly than any other group. They also tend to evaluate the need aggression more highly than do Americans

TABLE 2
EPPS NEED MEAN SCORES AND SOURCE OF SIGNIFICANT DIFFERENCES ON SOCIAL-DESIRABILITY RATINGS
BY AMERICANS, CANADIANS, HUNGARIANS, AND JUVENILE DELINQUENTS

Need	De- linquents (D)	Hun- garians (H)	Ca- nadians (C)	Amer- icans (A)	DH	DC	Differences*		HA	CA
							DA	HC		
Achievement	49.07	52.38	54.97	55.97	—	.01	.01	—	—	—
Order	48.36	60.60	53.01	53.81	.01	.01	.01	.01	.01	—
Exhibitionism	45.61	38.71	40.85	42.09	.01	.05	—	—	—	—
Autonomy	51.33	48.07	43.53	45.35	—	.01	.01	.05	—	—
Affiliation	58.95	56.36	60.85	59.75	—	—	—	.01	.05	—
Intracception	48.13	53.02	52.16	51.34	.01	.01	.05	—	—	—
Dominance	39.02	37.85	46.02	42.20	—	.05	—	.01	—	—
Change	60.44	51.70	53.61	54.31	.01	.01	.01	—	—	—
Endurance	48.78	57.97	56.58	53.88	.01	.01	.05	—	—	—
Heterosexuality	54.75	47.85	48.52	51.42	.01	.01	—	—	—	—
Aggression	48.87	42.68	32.19	30.99	—	.01	.01	.05	.05	—
Deference	47.34	47.21	51.49	50.98						
Succorance	44.72	48.12	47.91	50.66						
Nurturance	54.53	57.51	58.34	57.30						

* Entries show differences significant at the .05 and .01 levels.

and Canadians. Hungarians rate the need affiliation as being less socially desirable than the reference groups. This group of Hungarian students could be described as having acted in accord with the values they espouse. During a time of relative disorder in their own country they attained order by moving to another continent. Breaking away from their homeland may have been made easier by their relative underrating of the need affiliation. Of course they could also be rating order as relatively more socially desirable and affiliation as relatively less socially desirable as a later justification for their action.

Compared to American and Canadian college students, Canadian delinquents rate as more socially desirable the needs autonomy, change, and aggression. They also tend to rate exhibitionism and heterosexuality as more socially desirable than do the Canadians, but no significant differences appear between delinquents and Americans on these trait ratings. Delinquent mean rating of the needs achievement and order indicate they consider these traits less socially desirable than do the reference group. They tend also to undervalue the needs of intraception and endurance. Perhaps almost anyone who is forcibly institutionalized might be expected to value autonomy and change as socially desirable. On the other hand, the comparative low value placed on achievement and order corresponds to their particular deviant position in society. From these data it is impossible to say either (*a*) that the value system as reflected here has in part determined the delinquency or (*b*) that being in the role of institutionalized delinquent has determined the value system. It is possible to say here, as in the case of the Hungarian students, that the behavior and position of the group seem logically congruent with the need ratings.

To turn to another aspect of these data, Canadian delinquents differ more from American and Canadian university students than do the Hungarian university students, despite the fact that Hungarians come from a different subcultural background and language family. It may be that what has been called the social-desirability stereotype persists across widely different groups provided these groups are comparable in intelligence, education, and social class within the Western European cultural continuum.

Personality tests are usually constructed, standardized, and cross validated on groups available through institutions. By far most groups are students. It is tempting to assume that if groups which vary as widely as Americans, Canadians, Norwegians, Nisei, and Lebanese students indicate a stereotype of what is socially desirable behavior, then subgroups within one nationality must not differ significantly from this norm. However, farmers, loggers,

fishermen, and delinquents may have different ideas about what behavior is socially desirable. Individuals from these groups may produce statistically abnormal scores on a personality test, but they are quite likely responding in terms of the social-desirability norms of their own group. Similarities in value judgment within social class across national lines may be greater than across social-class lines within a nationality.

REFERENCES

1. DUNCAN, D. B. Multiple range and multiple F tests. *Biometrics*, 1955, **11**, 1-42.
2. EDWARDS, A. L. Manual for the Edwards Personal Preference Schedule. New York: Psychological Corp., 1954.
3. ———. The Social Desirability Variable in Personality Assessment and Research. New York: Dryden, 1957.
4. FUJITA, B. An investigation of the applicability of the Edwards Personal Preference Schedule to a cultural sub-group. Unpublished Master's thesis, University of Washington, Seattle, 1956.
5. KLETT, J. C. The stability of the social desirability scale values in the Edwards Personal Preference Schedule. *J. Consult. Psychol.*, 1957, **2**, 183-185.
6. KLETT, C. J., & YAUKEY, D. W. A cross cultural comparison of judgments of social desirability. *J. Soc. Psychol.*, 1959, **49**, 19-26.
7. KRUYTBOSCH, C. Flight and resettlement of the Sopron forestry faculty: A study of group integration and disintegration. Unpublished Master's thesis, University of British Columbia, Vancouver, B. C., 1958.

Department of Psychology
Western Washington State College
Bellingham, Washington 98225

THE EFFECTS OF DISCREPANT INFORMATION ON
PERCEPTUAL PROCESSES: AN APPLICATION OF
BALANCE THEORY*¹

Behavioral Science Project, University of California, Berkeley

DAVID VERN STIMPSON

A. INTRODUCTION

Balance theory, as set forth in *The Psychology of Interpersonal Relations* (2), is a general theory of cognitive balance that deals with the relationship between two people, P and O, and an impersonal object, X. Heider does not suggest operations by which to quantify these elemental relationships, although others (1) have been successful in doing this. The present study involved two modifications of Heider's original formulation: (a) incorporation of a quantitative indication of relative balance or imbalance and (b) alteration of the POX system to include three perceptual entities: an individual's self-perception, his perception of a group of which he is a member, and his perception of how the group perceives him.

In order to qualify elements that in Heider's original model are dichotomously positive or negative, they were assumed to fall on an evaluative continuum with the poles defined as high-positive evaluation and high-negative evaluation. The point along this continuum at which an element was placed by a subject was used as an indication of his perception of the value of that element. By assigning an ordinal scale to the continuum, the quantitative assessment of any particular element could be obtained.

The general notion of a strain toward balance indicates that if a balanced (harmonious) state does not exist, forces toward this state will arise (2, 4, 7). This balance hypothesis was utilized for making predictions of relationships among the quantitatively-assessed perceptual elements. It was assumed that a cognitive system made up of N elements would be balanced to the degree that the same numerical evaluation was assigned to all N elements and unbalanced to the degree that a disparity existed between the evaluations assigned to the N elements. A similar hypothesis, concerned with the relationship between

* Received in the Editorial Office, Provincetown, Massachusetts, on April 6, 1964. Copyright, 1965, by The Journal Press.

¹ Appreciation is due Kenneth R. Hardy and William G. Dyer, Brigham Young University, for making their students available and for administering the instruments used in this research.

two cognitive elements, has been experimentally supported by Osgood, Suci, and Tannenbaum (6).

In the present study, answers were sought to the following two questions: (a) What effect does the introduction of discrepant objective information into an individual's cognitive system have on the related elements of the system? and (b) Are there consistent differences among individuals in the amounts of cognitive discrepancy they will tolerate?

The above questions were translated into two formal hypotheses: Hypothesis I. If a discrepant objective element is introduced into a congruent subjective system, subjective elements will undergo re-evaluation in order to restore congruity to the system. Hypothesis II. There are relatively stable individual differences in the amount of incongruity that will be tolerated in any particular cognitive system.

B. METHOD

In order to test the hypotheses, we made direct comparisons of the evaluations assigned to several different objects by the same person. The semantic differential scale devised by Osgood, Suci, and Tannenbaum (6), which utilizes pairs of bipolar adjectives, was chosen to facilitate this comparison. Since these scales are relatively independent of the immediate conceptual context, they are well suited for the purposes of this study. Six pairs of adjectives were used to quantify the cognitive elements: good-bad, necessary-unnecessary, worthless-valuable, successful-unsuccessful, influential-uninfluential, and important-unimportant.

1. *Subjects*

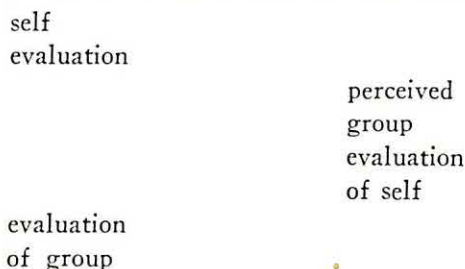
Subjects were 42 upper-division men and women students at Brigham Young University. They were enrolled in two sections of a class in group relations; 20 students were in one section, 22 were in the other. The sections met twice a week for two hours per meeting throughout the semester. The sections were conducted primarily according to group-centered training techniques, with no formal agenda.

Each subject was asked to rate four groups of objects on each of the six scales, with a maximum possible range of seven points on each scale: himself as a member of the group, the group, his perception of the group's evaluation of him, and each individual member of the group. A total score for each of the evaluated objects was derived by summing across the six scales, making a maximum possible evaluation of 42 and a minimum possible of seven for any one attitude object. Several pairs of distractor adjectives were included

in the scale but not scored; the direction of scoring was randomized to control for response set.

Students in both sections first responded to the semantic-differential scales after they had met together for a total of approximately 14 hours of interaction. From the responses it was determined how each person evaluated himself as a group member, how each person evaluated the group as a whole, how each person perceived the group as evaluating him, and how each person evaluated every other member of the group. By combining scores, the average evaluation given each individual by the group was also determined.

As was indicated earlier, the cognitive system of concern in this investigation consisted of three elements which may be diagrammed as follows:

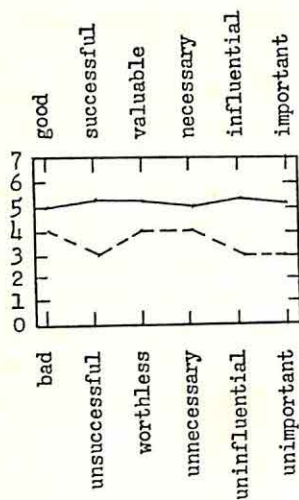


In keeping with the predictive scheme proposed, the cognitive system will be congruent to the degree that all elements of the system are similarly evaluated. It will be incongruent to the degree that there is lack of agreement among the numerical values assigned to the elements.

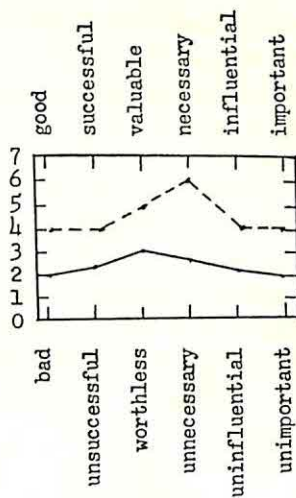
2. *Treatment*

Subjects were divided into three groups: (a) positive incongruent—when the average group evaluation of *S* was higher than his evaluation of himself, (b) congruent—when the group evaluation of *S* was approximately equal to *S*'s evaluation of himself, and (c) negative incongruent—when the group evaluation of *S* was lower than *S*'s evaluation of himself. Each subject was presented with a graph on which were indicated his self-evaluation and an average evaluation *purportedly* assigned to him by the group. This group evaluation had been changed (without *S*'s being aware of it) to increase the disparity between ratings for incongruent *S*s and to minimize the disparity for congruent *S*s. Sample graphs for each of the three conditions are given in Figure 1. A total of 14 students fell into each of the treatment conditions indicated in Figure 1.

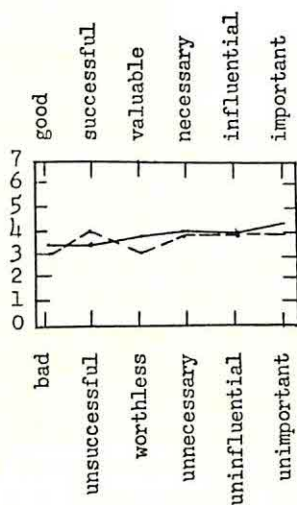
It was assumed that presentation of these graphs would create a situation



POSITIVE INCONGRUENT



NEGATIVE INCONGRUENT



CONGRUENT

--- Self evaluation of S
 — Purported evaluation of S by the group

FIGURE 1

SAMPLE GRAPHS ILLUSTRATING POSITIVE-INCONGRUENT FEEDBACK, NEGATIVE-INCONGRUENT FEEDBACK, AND CONGRUENT FEEDBACK

of cognitive imbalance in the incongruent conditions and a situation of cognitive balance in the congruent condition. The latter condition was used as a comparison group against which to test the effects of the experimental treatment.

The graphs, constituting the experimental treatment, were given to the subjects on approximately the third meeting of each section after the members had responded to the set of semantic differential rating scales. *Ss* were given 10 minutes to examine the graphs. Then each *S* indicated on a seven-point scale the amount of discrepancy he perceived between his self-ratings and the rating assigned to him by the group, as well as the amount of accuracy he attributed to each of the ratings. The graphs were then handed to the instructor, and the subjects again responded to the set of semantic differential scales.

C. RESULTS

To determine whether the "incongruent" situation created by the experimenter was so perceived by the subjects, a Spearman rank-difference correlation (8) was calculated between the amount of "incongruity" presented to each subject (Figure 1) and the degree to which he perceived incongruity to exist. The ρ is .81, is significant beyond the .001 level; therefore it seems safe to assume that the situation defined as incongruent was in fact phenomenologically incongruent.

Resolution of incongruity in a manner not covered by the experimental hypotheses could have occurred if a subject had rejected the accuracy of the objective discrepant information presented to him (3). To test this possibility a Wilcoxon matched-pairs signed-ranks test (8) was done on the perceived accuracy ratings. Results indicate no significant difference between positive and negative incongruent subjects in the amount of accuracy they attributed to the group's rating of them as compared with their self-rating.

Data relevant to Hypothesis I were analyzed by a Lindquist Type VI factorial analysis of variance with pseudoreplication (5).

The pretreatment and posttreatment evaluations constitute Factor A. The objects being evaluated (self, group, or perceived group rating of self) constitute Factor B. The condition of the subjects (positive incongruent, negative incongruent, and congruent) constitute Factor C. The results of the analysis are summarized in Table 1.

The significant treatment-by-subjects interaction indicates a difference between pretreatment and posttreatment evaluations assigned by subjects,

TABLE 1
ANALYSIS OF VARIANCE OF EVALUATION SCORES

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>
Between <i>Ss</i>	23.557	41		
Subjects	3.32	2	1.66	3.198
Error (b)	20.237	39	.519	
Within <i>Ss</i>	46.073	210		
Treatment	.803	1	.803	6.42*
Objects	16.715	2	8.357	49.159**
Interaction				
Treatment-by-objects	.099	2	.049	1.581
Treatment-by-subjects	5.608	2	2.804	22.43**
Objects-by-subjects	1.743	4	.438	2.576*
Treatment-by-objects- by-subjects	.575	4	.144	4.645**
Error (w)	20.530	195	.105	
Error ₁ (w)	4.882	39	.125	
Error ₂ (w)	13.235	78	.17	
Error ₃ (w)	2.413	78	.031	
Total	69.63	251		

* $p < .05$.

** $p < .01$.

which difference depends on the particular treatment given (*viz.*, positive incongruent, negative incongruent, or congruent). This significant difference results from a tendency on the part of positive incongruent subjects to raise their self-evaluations, group evaluations, and perceived group rating of self-evaluations; and for negative incongruent subjects to lower their evaluations. Evaluations of congruent subjects remained about the same. The tendency for the elements of a subjective system to undergo re-evaluation in order to establish congruity with a (previously) discrepant objective element confirms Hypothesis I.

The significant objects-by-subjects interaction indicates that evaluations varied as a function not only of the object being evaluated, but also of the treatment given the subject. Apparently subjects re-evaluated the three objects selectively, depending on the kind of incongruity they experienced. Table 2 shows mean posttreatment change in evaluations made by subjects in the three experimental conditions.

TABLE 2
MEAN CHANGE IN EVALUATION RESULTING FROM EXPERIMENTAL TREATMENT

Group	Self- evaluation	Evaluation of group	Perceived group eval- uation of self
Positive incongruent	5.07	3.78	4.93
Congruent	1.00	.14	2.14
Negative incongruent	-4.14	-.72	-3.38

A subject's hesitancy to lower his evaluations of the group even when the group rated him lower than he thought it would, but a willingness to raise his evaluations of the group when the group rated him higher than he thought it would, is an interesting phenomenon that deserves further investigation.

The higher-order interaction (treatment-by-objects-by-subjects) confounds replication variance with condition variance; so is uninterpretable.

Hypothesis II (namely, that there are relatively stable individual differences in the amount of incongruity that will be tolerated in any particular cognitive system) was tested by calculating rho between the total incongruity present in a cognitive system before and after the experimental treatment. It is .53; significant beyond the .001 level. Straightforward acceptance of the hypothesis is hindered, however, by further analysis of the data. Rhos computed separately for positive incongruent S_s , negative incongruent S_s , and congruent S_s are .63, .53, and .32 respectively. This lack of equivalence will be considered in the discussion section.

D. DISCUSSION

Our data support balance theory, extended to quantitatively-assessed perceptual elements in the cognitive system of individuals in a relatively enduring interpersonal situation. The fact that related elements of the cognitive system are re-evaluated by selective amounts to restore relative congruity when an incongruent objects element is introduced allows added insight into the dynamics of cognitive systems. The majority of investigators of the thesis of cognitive balance have focused on the element of the system that is most directly affected by the discrepant information. Our data indicate the importance of examining a broader spectrum of effects, including the effects on related elements of the system.

Our finding that the amount of change in evaluation of the group is dependent on whether a subject is in the positive or negative incongruent condition raises question as to the generality of the finding of "congruency" for both positive and negative conditions as reported by Taguiri, Bruner, and Blake (9). The data reported by Taguiri *et al.*, however, were based on person-person relationships; while our data were derived from person-group relationships. This difference may account for the lack of agreement between results.²

The fact that positive incongruent subjects show a higher tendency to return the cognitive system to a state of relative congruency than negative

² Another possibility is that there are personality differences between the positive and negative incongruent subjects in our study. Since no personality data were collected, this hypothesis must remain a speculation.

incongruent subjects is probably an artifact of the reluctance on the part of negative incongruent subjects to re-evaluate the group. This reluctance results in a larger discrepancy between the evaluation assigned the group and the other elements of the system by the negative incongruent subjects than was the case with the positive incongruent subjects.

Our results demonstrate (*a*) that balance theory is a useful predictive principle even in situations not saturated with demand characteristics; (*b*) that the introduction of discrepant information into a cognitive system causes reverberations and adjustments through an extensive part of the affected cognitive area; and (*c*) that different cognitive dynamics operate under conditions of positive and negative incongruence.

E. SUMMARY

Our purpose in this study was to test a quantitative extension of Heider's balance theory in an extraexperimental setting. We sought answers to two questions: Does the introduction of discrepant objective information into an individual's cognitive system have an effect on related elements of the system? and Are there consistent individual differences in the amount of cognitive discrepancy that will be tolerated?

Heider's balance formulation was quantitatively extended by having subjects rate relevant attitude objects on a semantic-differential scale. Each subject rated himself, each member of the group, the group as a whole, and his perception of the group's rating of him. Two weeks later, subjects were given feedback in the form of individual graphs that indicated the rating they gave themselves and the purported (but distorted) average rating given them by the group. Then they rated the same attitude objects on the same scales again.

Comparison of each subject's initial ratings with his final ratings indicated the degree to which the feedback information was incorporated into his cognitive system and the degree to which perceptions were changed in order to achieve balance among the elements of the system.

Results indicate that related elements of the cognitive system were re-evaluated by selective amounts in order to restore relative balance to the system following the introduction of objective discrepant information.

Subjects who received positive feedback (group rating of *S* higher than *S*'s rating of himself) showed a greater tendency to restore the system to balance than subjects who received negative feedback (group rating of *S* lower than *S*'s rating of himself). It was impossible to discover whether this was due to situational variables or personality variables because the subjects in the two experimental conditions were self-selected.

REFERENCES

1. CARTWRIGHT, D., & HARARY, F. Structural balance: A generalization of Heider's theory. *Psychol. Rev.*, 1956, **63**, 277-293.
2. HEIDER, F. *The Psychology of Interpersonal Relations*. New York: Wiley, 1958.
3. HOVLAND, C. I. Reconciling conflicting results derived from experimental and survey studies of attitude change. *Amer. Psychologist*, 1959, **14**, 8-17.
4. JORDAN, N. Behavioral forces that are a function of attitudes and of cognitive organization. *Hum. Relat.*, 1953, **6**, 273-287.
5. LINDQUIST, E. F. *Design and Analysis of Experiments in Psychology and Education*. Boston: Houghton, 1956.
6. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957.
7. ROSENBERG, M. J., HOVLAND, C. I., MCGUIRE, W. J., ABELSON, R. P., & BREHM, J. W. *Attitude Organization and Change*. New Haven: Yale Univ. Press, 1960.
8. SIEGEL, S. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.
9. TAGUIRI, R., BRUNER, JR., & BLAKE, R. On the relation between feelings and perception of feelings among members of small groups. In Maccoby, E., Newcomb, T., & Hartley, E. (Eds.) *Readings in Social Psychology*. New York: Henry Holt, 1958.

265 East 1st South
Provo, Utah

GROUP RESPONSES TO THREAT: PERCEPTION AND LEARNING UNDER STRESS*¹

Department of Psychology, Rice University

BRADFORD B. HUDSON, JOHN McDAVID, AND MARIAN ROCO BINNER

A. INTRODUCTION

There are many human behaviors that could be studied to better advantage under field conditions than in the *as-if* laboratory environment. Many extreme situations, such as those encountered in disasters cannot be duplicated in the laboratory, as Guetzkow (9) has pointed out; but there are some that can be approximated. Those that simulate life conditions are useful in the study of emotional behavior in the light of its dependence upon spontaneity.

In this report, the authors are concerned with experiments in which spontaneous emotional responses were elicited in groups by ambiguous environmental events that were perceived as threatening. As in life situations, neither the perceptions nor the responses were stereotyped. Preliminary studies revealed in fact a rich variety. Accordingly, the unique response as well as the uniformities in responses are reviewed. Earlier studies with animals have indicated the generality of avoidance responses to strange or unfamiliar stimuli. These studies with humans, an extension of this earlier work, imply that on the emotional level qualitatively similar behaviors are involved.

In discussing the responses elicited in these studies, the term *anxiety* has been employed to characterize the feeling reported by many of the subjects. Murray (22) has used it to designate a broader class of response to include apprehension and worry. In situations more extreme than the ones studied here other terms might be used: e.g., terror or panic. Agreement, however, is not high as to the meaning of this whole class of terms; for example, Webster's dictionary refers to panic as a sudden and overpowering fright and ultimately lists fear as a synonym. Irving Janis (16), Joost Meerloo (20), and Grinker and Spiegel (8) have treated panic as behavior; while psychoanalytic theory tends to treat panic as a subjective feeling. It has been discussed by

* Received in the Editorial Office, Provincetown, Massachusetts, on April 6, 1964. Copyright, 1965, by The Journal Press.

¹ These studies were supported by the Office of Naval Research Nonr-490(01) and the Disaster Committee of the National Research Council. The authors are grateful for the assistance of Dr. James McCary of the University of Houston, Officers of Bergstrom Air Force Base, and the Fire Department of the City of Houston.

Harry Stack Sullivan (25) and Hadley Cantril (5) as a subjective state tending to lead to relatively consistent and predictable modes of behavior. Other terms, such as fear or apprehension, might serve as well (21); however since the stimuli to which the subjects in these experiments were exposed was characterized by ambiguity of instrumental origin or of meaning the term anxiety seems appropriate in this context—a usage not necessarily coinciding with the clinical concept.

Several experiments conducted in different settings and with subjects varying in age and experience provide the background for this report. The first experiment was exploratory—the experimental situation was a classroom and the subjects were Rice University students in a course in introductory psychology. The second experiment was conducted in the same setting during a regularly scheduled class, but three years later. In addition to these two studies, three replications conducted with small groups of adults in business or professions attending a two-week course in civil defense provided additional qualitative data.

B. METHOD

Subjects in the main experiment were 62 undergraduate students at Rice University in a class at which approximately 100 were in attendance on the experimental day. Of these, 31 males and 31 females who were available for appointments and interviews during the week following, comprised the sample.² The basic procedure was as follows. During a regular lecture, a sequence of sounds was presented. These sounds, assembled on a prepared tape, were broadcast from a series of loudspeakers surrounding the building in which the lecture was in progress. The sounds were presented over a period of approximately 25 minutes, beginning a few moments after the class started. The first sounds were those of newsboys shouting "Extra" at a distance beyond distinct audibility. After several minutes, there were sounds of fire equipment and again after intervals sounds of distant aircraft. Gradually, the amount of aerial and ground activity developed to a point where intervals between events represented by the sounds of sirens and mass flights of aircraft, conventional and jet, were short. Toward the end of the 25-minute period, several people were seen running past the open doorway of the lecture hall, and at about the same time a stranger entered the room to deliver a note to the lecturer. The climax came at this point, accompanied by an air-raid alarm,

² Not included in the sample were those students who were aware that an experiment was being conducted. This group consisted of those who knew of the experiment conducted three years prior and those who spotted the moving-picture camera concealed in the lecture room.

and almost simultaneously by the sounds of distant explosions and a flash of light outside the building.

The buildup of events was gradual and not (until toward the end) intrusive. The subjects were provided no real basis for rejecting assumptions that the events that developed during the 25-minute period were real. If they accepted as real the presence of newsboys calling "Extra" at the beginning of the hour, it was difficult to reject the evidence of subsequent events which provided that initial experience with meaning. On the other hand, the course of events portrayed by the tape recording was such as to preserve a high degree of ambiguity until perhaps the last few seconds, with the sounds of distant explosions. The variety of sounds and the rapid development of the situation provided no satisfactory basis for the formulation of hypotheses that logically would account for *all* of the activity evidently taking place.

Upon the termination of the sound sequence, an experimenter entered the lecture hall, advised the group that they had been participating in an experiment, and asked that they remain seated for the administration of a questionnaire. The questionnaire contained 34 items on reactions to and interpretations of the experimental situation. Two days after the first questionnaire a second was administered. During the following days the subjects were met for personal interviews at which time, in addition to completing specific follow-up questionnaires, they were encouraged to give detailed unstructured accounts of reactions during the experimental period.

Observational data were recorded by four observers planted in the lecture hall for both the two preceding lectures and on the experimental day. Each recorded the behavior of selected groups of students, observing each subject in turn for periods of 10 seconds during the period of the experiment. A record was made of conversation with others, posture, and gross body-position changes. From observer records, considerably more "looking around" was evident on the experimental day, the students attempting to catch glimpses through the high windows of the aerial activity evidently in progress outside; there was more conversation, but the observers were unable to detect other qualitative differences in behaviors on the experimental day. They did not report apparent signs of anxiety or of emotional reactions.

In both the pilot and the main experiment, moving pictures were taken from a concealed position, the field of the camera encompassing some 18 students. The photographic data corroborated the observers' record in that the moving pictures failed to reveal any individual differences systematically associated with measures of anxiety subsequently determined from questionnaires and interviews.

C. RESULTS AND DISCUSSION

The strength of the anxiety response was the major dependent variable in these studies, and of primary interest are the relationships that other measures bear to that response. On the questionnaire administered immediately after the termination of the experiment, each subject was asked to estimate graphically the development and the amount of anxiety he experienced during the course of the experiment. Figure 1 is a sample record of a student's response. The high point on the curve was accepted as the individual's score on self-rated anxiety and was one component in the measure of anxiety.

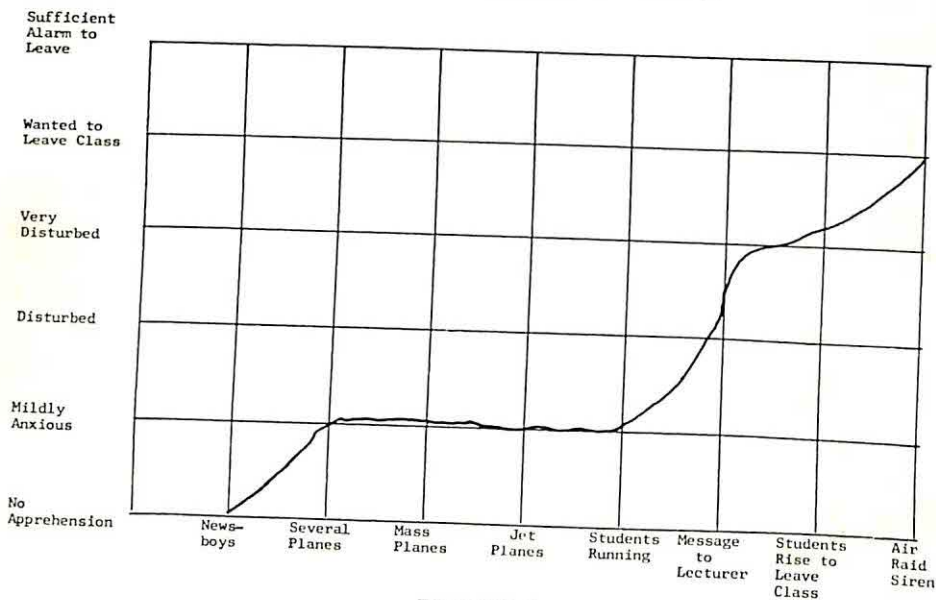


FIGURE 1

SAMPLE GRAPH OF SELF-RATED ANXIETY

(The directions were as follows: "Plot the development of anxiety or apprehension that you experienced. . . ." Events are indicated on the base line and degree of anxiety is indicated on the vertical axis.)

This score and data from six other questions³ provided the basis for the

³ Responses to the following questions provided the data from which the judged anxiety score for each individual was derived: (a) Did you experience any aftereffects of the experiment? (b) Did sounds of sirens or planes after the experiment ever remind you of the experiment? Were there any attendant feelings of anxiety? (c) Did you have any feeling that you or anyone else was in personal danger during the experiment? (d) When you learned that it was an experiment how did you feel? Relief? Anger? Interest in results? Disappointment? Unpleasant letdown? Any other? (e) Did any plan or plans of action flash through your mind as the experiment progressed? (f) What experience, if any, in your life has produced apprehension

individual's measure of anxiety. This score was determined by a panel of four judges, each independently sorting the subject's records into one of six categories on the basis of the seven questions. Disagreement among judges on score assignments was discussed until agreement was achieved. Table 1

TABLE 1
FREQUENCIES IN ANXIETY CATEGORIES

Response magnitude	Judged rating		Self rating	
	Males	Females	Males	Females
"disturbed to the point of leaving the class"	3	3	2	5
"disturbed to the point of wanting to leave class"	5	6	10	9
"very disturbed"	7	12	8	5
"disturbed"	9	9	3	10
"mildly anxious"	5	1	5	2
"no anxiety"	2	0	3	0
Total	31	31	31	31

represents the frequencies in the several categories for men and women and shows both the judged ratings based upon the responses to the seven questions listed and the self-rating, as determined from the graphs of anxiety levels illustrated in Figure 1. Distributions of the measures of anxiety show slightly less dispersion (with fewer cases in the extreme categories) than do the self-rating scores.

The results are treated from two points of view. The first, focused primarily on the main experiment, concerns the quantitative relationship between the strength of the anxiety response and other measures employed. The second viewpoint concerns qualitative descriptions. It was evident from responses to the questionnaires and from interviews that some of the data would not lend themselves to quantitative treatment. The perceptual structuring of the situations directly or indirectly reflects the past experience of each individual; hence the variety of individual reactions, which constitute an important part of the data, are treated separately.

The data obtained from the main experiment were analyzed in terms of a set of 25 variables derived from content analysis of the interview and questionnaire records. Since data on a substantial number of variables were categorized, the contingency coefficient, C , was used as a measure of association. The significant interrelationships observed among variables are listed in Table 2.

comparable to that developed during the experiment? Please describe fully. (g) Plot the development of anxiety or apprehension that you experienced on the chart below. Events are indicated on the baseline, and degrees of anxiety are indicated on the vertical axis.

TABLE 2
SIGNIFICANT RELATIONSHIPS AMONG EXPERIMENTAL VARIABLES

Variables	Contingency coefficient ^a	Significance level ^b	Approximate Pearson r^c
Anxiety (with)			
Hypothesis consistency	.45	.05	— .57
Number of alternative hypotheses	.47	.01	.60
Estimated (projected) group anxiety	.53	.01	.71
Estimated (projected) group anxiety (estimate graphed)	.63	.01	.72 ^d
Suggestibility	.31	.10	.41
Tendency to "scapegoat"	.36	.05	.48
Aftereffects	.33	.05	.45
Retention	.31	.10	— .42
Hypothesis consistency (with)			— .74 ^e
Suggestibility	.34	.02	— .48
Specificity of perceived threat	.32	.05	.45

^a Derived from chi square.

^b From value of chi square.

^c C corrected for categorization.

^d Computed Pearson r .

^e Computed rho for special group [see text, p. 130].

The experimental procedure was uniformly effective in producing anxiety. The average strength of the response varied from group to group for reasons that are complex. Part of the variability reflects the type of interpretations or hypothesis formulated by individual subjects. These in turn are in part a function of chance factors; for example, in the pilot study the fact that the Texas City disaster was a well-remembered event tended to influence the interpretations made by some of the subjects. Tensions that happened to exist at the time between the United States and the Soviet Union were reflected in the frequency of interpretations of war. Interpretations, in other words, were often associated with recent experience. This process of hypothesis formation or interpretation of events is the first subject for discussion in the paragraphs below.

1. Hypothesis Formation

With few exceptions, individuals exposed to the experimental situations attempted to organize the events in meaningful ways. This necessarily involved relating them directly or indirectly to past experience. This process involved the formulation of hypotheses or tentative interpretations in which war, accidents, and fire were common. As noted, the influence of current or recent events was frequently evident; for example, "The morning newscast's discussion of peace made the possibility of war seem less likely." Table 3

TABLE 3
TYPES OF INTERPRETATION

Anxiety level	War %	Accident %	Fire %	Experiment %	Other %	Total %
5	50	50	0	0	0	100
4	46	9	27	0	18	100
3	11	47	16	0	26	100
2	22	11	0	17	50	100
1	0	50	0	0	50	100

shows the per cents of the several interpretations for each level of anxiety in terms of the dominant hypothesis held by the subject.

Two measures related to hypothesis formation were employed. The first was an index of the stability of an hypothesis and the consistency of the individual in his structuring of the situation. This measure was based upon answers to several questions: (a) the number of hypotheses, (b) the frequency with which the subject changed his interpretation, (c) the extent to which he was influenced by others, and (d) the degree of consistency between the hypotheses and the events portrayed during the experiment. To convert the several responses into a single score for each subject, each questionnaire was rated independently by four judges on a scale ranging from 1 to 6. Discrepancies in ratings were adjusted in discussion among the judges. The 62 subjects were grouped as indicated in Table 4.

TABLE 4
HYPOTHESIS CONSISTENCY

Anxiety level	N	Good consistency	Moderate consistency	Inconsistent, disorganized
4-5	17	0	7	10
3	19	3	9	7
2	18	5	10	3
0-1	8	4	2	2
Total	62	12	28	22

The relationship between hypothesis consistency and anxiety is negative [$C = .57$ ($p < .05$)]. In other words, the greater the individual's success in developing a structured, internally consistent explanation of events the lower his level of anxiety.

Hypothesis consistency is positively related to a second measure, perceived specificity of threat [$C = .32$ ($p < .05$)]. The score on this measure was based upon responses to three questions: (a) Did a plan or plans of action flash through your mind as the experiment progressed? Describe. (b) Assuming that this had been a real disaster, on leaving class what would you have done

first? where would you have gone? what specific route would you have taken?
(c) What thoughts of the immediate or distant future flashed through your mind as the experiment progressed?

One element in the measure of hypothesis consistency was the number of alternative hypotheses. To obtain an estimate of its relationship to anxiety, it was treated separately as well. The subjects were classified into three groups according to the number of hypotheses reported: Group 1 (one hypothesis), $N = 2$; Group 2 (two to four hypotheses), $N = 44$; Group 3 (five or more hypotheses), $N = 16$. A C of .47 ($p < .01$) indicates that the fewer the number of hypotheses the lower the level of anxiety. These relationships are interesting in the light of another set of results that pertain to efforts by the subject to verify hypotheses.

2. *The Search for Information*

Persistent attempts to verify hypotheses were characteristic in each replication of the experiment. In the classroom situation, sources of reliable information were of course nonexistent. There were, however, three pseudosources: communication with individuals in the class, observations of the lecturer by the class, and observations of the class as a whole.

The tendency of subjects to accept tentatively one hypothesis as the most likely even though incompatible with other features of the situation implies that under stress any belief that accounts for part or parts of the unknown may be uncritically accepted. This state of mind sets the stage for the ready acceptance of rumors and underscores the second significant behavioral characteristic observed, the search for information.

3. *Communication with Individuals*

As noted earlier, on experimental days planted observers recorded an increase in conversation among persons seated together in the classroom. Figure 2 indicates the nature of exchanges between a sample of subjects seated together, a group within the field of view of the moving-picture camera. The following exchanges are typical of the interplay of opinions: "My neighbor said it was war and this increased my anxiety," or "My neighbor said it was maneuvers from Ellington Field and this calmed me somewhat." These ideas and opinions clearly filled a need. They aided in providing some structure to an ambiguous situation, serving to confirm or disconfirm existing hypotheses. The fact that they could be nothing more than opinions as to what was happening seemed not to prevent such conjectures from contributing to the individual's increase or decrease of tension.

Seat number	114	115	116	117	118
Self-rated apprehension	Absent				
Remarks: About self-		2	3	2	3
From neighbor on left-		"What is going on outside"	"I was very curious"	"I wondered what it was"	"I thought of Pearl Harbor"
From neighbor on right-		"What the hell is happening"	"We were joking, trying to guess what was happening"	"He was very disturbed"	"She was very worried"
From neighbor in front-		Absent	"Do not recall"	"We wondered"	"Thought of an accident"
Self-rated apprehension	94	95	96	97	98
Remarks: About self-	4	4	3	3	4
From neighbor on left-	"In a complete state of confusion"	"Russia has declared war"	"I could not decide anything"	"I was extremely anxious. I started praying"	"I thought we had declared war with Russia"
From neighbor on right-	"He was as confused as I was"	"He seemed unconcerned"	"No comment"	"She kept saying, 'Dear God, No.'"	"Said nothing"
From neighbor in front-	Absent	"She thought of war"	"No comment"	"The world is falling apart"	"She was very disturbed. This scared me."
	"She did not say a word"	"He seemed apprehensive"	"No comment"	"He said nothing"	"Do not know"

FIGURE 2

SUBJECT INTERACTION

(A block of nine students seated within the field of the motion-picture camera. The quotations within each block are remarks or observations reported by the student in the indicated seat. Rating of apprehension: 4—wanted to leave class, 3—very disturbed, 2—somewhat disturbed, 1—mildly apprehensive, 0—no apprehension.)

The readiness with which these opinions and ideas were accepted and responded to emotionally illustrates the ease with which rumors may be propagated. Suggestibility was evident in remarks, such as "The neighbor in front of me said she saw a strange-looking plane flying around. My anxiety increased," "My neighbor saw something round in the sky through the window. I replied that maybe they are testing a new jet." These also illustrate the lowering of one's critical ability [see Cantril (4)], for in this situation in the absence of outside contact it was evident that no one was any better informed than anyone else. The intense desire for information was illustrated by the dead silence that fell over the class when toward the end of the sound sequence a person entered the room to deliver a note to the lecturer. For the class as a whole this was the point of maximum anxiety: The importance of the events was being "confirmed," but *what* was being confirmed remained unknown.

4. *The Nominal Leader*

The second pseudosource of information was the lecturer. In the search for a clue as to the meaning and portent of events taking place, he became a focus of attention. It was evident that he was observed intensely for evidence of anxiety, and that his failure to show it contributed to its reduction within the class. "The fact that the professor was not anxious reassured me," was a frequent remark in the subsequent interviews. The fact that the lecturer clearly had no access to outside information and that his sources were no better than those of the subjects themselves did not interfere with his evident composure serving a source of reassurance, another example of a lowering of the individual's critical ability.

Suggestive, although not statistically significant, is the degree of dependence upon leadership as a function of anxiety. Those Ss rated highest on anxiety (Group I) tended to rank "leadership" of first importance as a means of preventing panic in the sort of situation to which they had just been exposed (50 per cent). To be given information was listed second by this group (33 per cent). Among subjects rated second highest on anxiety (Group II) the order of importance was reversed, information being considered as most important and leadership second (55 per cent *vs.* 18 per cent). Those rated lowest on anxiety ranked other means, such as "To have a friend to rely upon," or "Each person having a job to do" (65 per cent), leadership dropping to six per cent.

The reliance, however, upon a person in authority for reassurance is an interesting illustration of the development within a group under stress of

attitudes of dependence upon a leader and of the leader's potential power in such an unstructured situation. One *S* remarked that when the lecturer left the room, "I felt that our leader had gone—an unpleasant feeling."

The fact that there were few instances of overt behavior as long as the lecturer remained in the room reflected in part the group's deference to him. There were a few exceptions. One student in a replication of the experiment left the lecture hall and, when questioned as to where he was going, sheepishly stated a desire to see if he could spot his brother's plane flying from the nearby air-force base. Another subject jumped for the shelter under a table at the moment of the flash outside the building, reflecting his interpretation of war and an already formulated plan of action.

5. *Projection*

The third source from which the subjects attempted to secure information as to the meaning or portent of events was the class as a whole. Here an interesting phenomenon routinely appeared, two different questions yielding essentially the same results. The first question, "Did you notice any uneasiness in the class as a whole during the experiment?" required responses in terms of "much," "some," or "none." A marked tendency was evident for the subject's observation of group apprehension to increase with the level of his own anxiety. The term projection has been used to designate this process, meaning (in other words) the perception of others as experiencing feelings and making interpretations similar to one's own. The association between estimated group anxiety and the measure of anxiety is .53 ($p < .01$).

The subjects were required to estimate graphically the development of anxiety over time in the class as a whole as they had for themselves [see Figure 1]. The correlation of this graphic measure of group anxiety with the graphic estimate of self-anxiety yields a Pearson r of .72, and with the scale measure of anxiety a C of .63 ($p < .01$). In other words, in terms of both measures of group anxiety, those subjects who were apprehensive tended to believe that they observed uneasiness within the class as a whole. Those who were not anxious perceived it less.

Consistent with these results are the facts that neither the nonanxious lecturer nor planted observers within the class during the experiment were able to detect evidence of anxiety. Further, careful study of the moving-picture record provided no clues that differentiated the most anxious from the least anxious subject. Since the individual subject was able to exchange opinions only with his immediate neighbors his observations were to some extent at least a projection of his own feelings. This interpretation is supported

by the subjects' reports of what they had perceived as signs of anxiety within the class. For example, the hilarious amusement of a small group of students who had seen the moving-picture camera and were aware that they were the objects of an experiment did not interfere with the development of anxiety in other persons seated elsewhere in the lecture room. Instead, their uproarious behavior was perceived as evidence of their anxiety.

Humor in the form of laughing, joking, or jesting about events occurring was reported by a number of subjects as evidence of the disturbed state of those displaying such behavior. In response to the question, "What were the signs of apprehension you observed in the class as a whole?" 42 of the 57 subjects who answered this question reported either restlessness (18 subjects), humor (13 subjects), or curiosity (11 subjects). Other answers were suspense and looking around. In other words, there is a wide range of behaviors which under stress can be accepted as confirming one's anxiety.

The average level of anxiety perceived in the class as a whole was higher than the average level of anxiety reported by the individual. This was evident from measures derived from the graphs drawn to represent the course of development of anxiety in self and that perceived for the class as a whole. It was also evident from other measures. This may be a symptom of an unwillingness in the individual to admit to the full extent of his anxiety or it may reflect a common hypothesis, "Maybe they know something that I don't." In any case this projection of one's own anxiety may be significant in group situations, for in perceiving anxiety in others one finds confirmation or overconfirmation for one's own undefined fears; thus tending to set up a cyclical reaction partially self-sustaining and possibly to some extent self-amplifying. In fact, this process may be a factor underlying the speed with which some panic reactions develop in groups. Furthermore, the fact that none of the methods of direct observation in either of the Rice studies revealed overt manifestations of anxiety within the class as a whole suggests another possible reason why panic in certain circumstances may break rapidly. Inhibition of emotional expression in our culture is socially approved. However, when a percentage of a group breaks over into overt activity, social sanctions are lowered, with the result that those who had been inhibiting expressions are rapidly recruited by the reacting group.

6. *The Nature of the Threat*

From the foregoing it is evident that the inability to organize events in one's environment perceptually in a stable and meaningful way is a factor underlying anxiety. Response to the strange or unfamiliar has been considered

by Hebb (11), in his observation of the reactions of chimpanzees to strangers and strange objects; and by Hudson (13), in studies of rat behavior where the typical rodent defense reaction of covering a strange object with shavings is evoked. Response to the unfamiliar has been discussed by Bruner (2), in his consideration of perceptual learning; and by Cantril (5), in terms of the organism's attempt to come to terms with its environment: i.e., to achieve clarity.

In this experimental situation the combination of events carried the implication that something important was happening. The newsboys on the campus, a rare occurrence, set the stage for this interpretation. However the major hypothesis of many of those who were disturbed did not explicitly incorporate this feature, "something important." The hypothesis that students at the neighboring university were advertising their "Fiesta Day" did not involve it, nor did the hypothesis of fire or accident really qualify. Further, none of these would have been announced by a special edition of the newspaper. The absence of logic in interpretation fostered exploration and search for better alternatives, a trial and check or hypothesis type of behavior.

It was also evident that threat *per se* was not necessarily involved at the conscious level. Although war-related hypotheses were most frequent among those who were most anxious, the hypothesis of fire or accident (neither of which directly involved threat) were often expressed by those in the high-anxiety categories.

It is an interesting question as to whether the gradual increase in average level of anxiety that continued to the end of this experiment was due (a) to repeated failures of hypotheses to remove the ambiguity, reducing events perceived from the strange to the familiar; (b) to the steady addition of new elements in the sequence of events; or (c) to both. Adaptation to the strange and novel may occur quite rapidly (13). The alerting process, a mechanism of a most basic sort, has served its biological purpose when it provides the organism with the stimulus to explore, to identify the strange, or to determine its implications. If there ensues a period of steady state after this initial alert, if there are no changes in the perceived character of events, the adaptation process may set in whether or not the stimulus has been recognized or organized cognitively into something meaningful. During the civil war in Spain, populations responded at first to the bombing of urban areas with great anxiety, taking refuge in available shelters; but with the repetition of aerial attacks they came outside to watch. Two men en route to England from the United States during World War II were among those rescued from a torpedoed vessel. They reported intense anxiety in response to the emergency.

Within hours the ship that had rescued them was torpedoed. Their reaction to the second experience, though it was equally fraught with danger, was reported as quite different. In this instance, they presumably had not only partially adapted to this type of situation but, by virtue of the first torpedoing, had resolved some of the ambiguities of action to be taken on the second occasion.

The hypotheses formulated during the period of uncertainty may be associated with a physical exploration of the stimulus or with checks on the logic or internal consistency of interpretations. These activities are highly motivated. It is as if to be without an organized view of the environment is intolerable; however the response needs to be understood in terms of complex intellectual process only insofar as recognition requires such processes. The rustle of leaves or the snap of twigs will alert an animal in the woods until the ambiguity of its origin is resolved, but the animal does not respond to signs of more complex events that alert a human adult. Similarly, children accept without anxiety events that are novel in terms of their experience for their failure to perceive the ambiguities involved, whether they be in terms of cause, implication, or outcome (18).

It is interesting, and it needs to be stressed, that it is not necessarily a threat component contained within an hypothesis that arouses anxiety. Hypotheses may even be such that if correct they would remove the element of threat or of danger which at first would appear to be the origin of anxiety. The hypothesis that "It is all an experiment" is an example, but that such an hypothesis by itself is not sufficient (unless confirmed) is evident in the persisting anxiety among those who held it.

Bruner (2) cites Bogoras, who reports that a Chukchee woman almost fainted on finding that he had brought a strange box into her house, objecting on the grounds of its strong and malevolent odor. However, instead of inferring that she behaved so because her hypothesis was that strange things smell bad, in terms of these data her emotional response reflected the absence of a perceptual organization to give meaning to this unfamiliar object with its strange odor.

7. *The Associative Process*

Hypothesis formation, a more primitive process than hypothesis behavior described by Kreshevesky (19), resembling in some respect that examined by Bruner (2), appears as a high-speed associative or "scanning" process for which the rules of frequency and recency do not seem to apply. For example,

in a series of experiments with small groups, in which ambiguous sounds were presented under conditions which appeared to represent real events in the environment, as many as 10 hypotheses have been recalled by subjects after exposure to ambiguous sound stimuli of 15-second to 30-second durations. In one instance, the stimulus was a recording of the sound of horses trotting on pavement, although never identified as such. The sounds came from a corner of a large laboratory which was empty except for two subjects present ostensibly for the purpose of grading papers. At one point one said to the other with overtones of alarm, "What is that?" to which the other responded, "Maybe we'd better get out of here."

The series of hypotheses rapidly formulated, checked, and rejected by one subject was interesting, reflecting associated past experiences of both a general and a specific nature. The generalized ones were "I thought of carpenters hammering on the floor. Then I thought of students coming down the stairs. At some point the idea of repair work on the ventilating system occurred to me and then I thought of the architects, but the sounds were coming from the wrong direction." One specific hypothesis was significant. This subject recalled a momentary scene in a moving picture attended some five years previously when she was 12 or 13 years of age—a locomotive was emerging from a tunnel. It suggests that some element in this experimental situation was common to this earlier flash experience in the moving-picture theater.

To account for this specific association, since it was not supported by recency or frequency, and it is questionable whether vividness *per se* was sufficient to account for its recall, it would appear necessary to postulate a one-trial exposure and the principle of identical elements (26) in which specific components in the immediate present initiate momentary associative elaborations starting with an element or elements common in the earlier experience in the theater. The specificity of this type of association should vary inversely with the frequency with which the identical elements have been experienced in the intervening period. Among adults, one would not expect a high frequency of such specific hypotheses, but among children for whom there may be many stimuli that still have a one-trial association, specifically elaborated hypotheses of the type described should be more frequent.

It will be recalled that in these experiments with large groups the more specific the hypothesis the lower the level of anxiety. Children in their spontaneous efforts to organize events may not only formulate more specific hypotheses, one-trial learning based; but in evaluating them accept them less critically for the lack of other incongruent rejection-forcing experiences.

8. *Aftereffects*

It had been observed in the exploratory study that for periods in some cases for several days, after the experiment, the sounds of planes or fire trucks would partially reinstate emotional feelings experienced on the experimental day. Therefore, following the main experiment the subjects were questioned during interviews as to the associations evoked during subsequent days. Of the 61 subjects answering these questions, 43 indicated the experience of such aftereffects; 18 did not. "I laughed with relief when I heard planes or sirens for several days after the experiment." "Whenever I heard an airplane the rest of the day I felt a frightened feeling."

The experience or nonexperience of these aftereffects is related to the strength of anxiety response [$C = .34$ ($p < .05$)]. Since there was a spurious element in this relationship, the anxiety measure being based in part upon evidence of such aftereffects, the latter is also correlated with the individual's uncorrected graphic self-rated anxiety, yielding a C of $.33$ ($p < .05$).

The immediate aftereffects reported on learning that it had been an experiment were varied. A number reported an unpleasant letdown; others, feelings of both relief and disappointment: "I was really wanting something to happen."

Another kind of aftereffect was frequently evident: "I felt let down *and* *peeved*,"—an expression of hostility toward persons associated with the experiment. It was sufficiently frequent (11 of 62 subjects reporting this feeling) that its relationship to the anxiety measure was examined. Dichotomizing in terms of presence or absence of this "scapegoating tendency" and relating it to anxiety yielded a C of $.36$ ($p < .05$).

9. *Efficiency of Recall*

With the availability of a moving-picture record of a small segment of the class, a sample of 10 subjects clearly and constantly within the field of view of the camera was studied in detail and a relatively complete record made of the behaviors of each subject during the experimental period. From these data, checklists of overt behavioral acts were prepared.⁴ These items were then buried within a list of irrelevant items and presented to each subject for recognition of his own overt behaviors. The scores, in terms of percentage of items recognized, for the 10 subjects were rank ordered and correlated with the anxiety scores, yielding a ρ of $-.74$ ($p < .01$). Supporting these

⁴ This special analysis was done by Jan Price and Judy Garland at Rice University.

results, the subjects in the class as a whole were also questioned as to whether they were aware of memory lapses during the experiment. Based upon these verbal reports, a C of .31 ($p < .10$) was obtained.

It is evident that under stress the field of observation narrows, reducing self-awareness to varying extents. In terms of the verbal reports, the subjects became so involved in the process of attempting to interpret the situation that there were numerous instances of feeling after the experiment that there were time gaps, a lack of complete continuity in the recall of events. These results perhaps can be related also to Ranson's suggestion (24) that "Panic . . . involves temporary major disorganization of thinking and control of fear. Consciousness is usually clouded."

10. *Other Behaviors*

This type of situation leads to behaviors that, though they do not permit statistical treatment, are interesting in demonstrating the range of response patterns that may emerge under stress. For example, disassociated physiological reactions have been reported. One subject stated in the interview that he had not been aware of being disturbed or anxious during the experiment, but he admitted to feelings of hostility toward those responsible. With probing it was also brought out that he had perspired noticeably more than usual during the period of exposure to the sounds. Another subject, somewhat puzzled by her failure to experience the anxiety that she had detected in her neighbors, recalled that her hands had become very moist, something she had not noticed until she started to fill out the questionnaire.

A question, "What did you do or say to yourself to reduce your apprehension?" called forth a variety of replies. "I started talking to my neighbor to forget about what I thought was happening." Another, "We are all in the same boat"; a girl, "I said a short prayer and after that I did not worry anymore"; a boy, reporting the statement of a neighbor, said, "Glad I went to church."

Humor as noted, was often evident: "I shot the bull," or, when the lecturer left the room, "Maybe the professor has been drafted." or "Next comes the atomic bomb," and "What's that? It sounds like space ships."

11. *Ambiguity in Relation to Extreme Behaviors*

In discussing panic behavior Caldwell, Ranson, and Sachs (3) list at least six factors that contribute to its arousal. Among them are (a) tension and insecurity, (b) imitative behavior, (c) suggestibility, (d) rumor, (e) fear of the unknown, and (f) lack of preparation. In terms of these studies, insecurity

and tension would appear as aspects of the anxiety response; and responsiveness to rumor, a symptom of the need for organization (the lack of which underlies anxiety). Equivalent is suggestibility and similarly symptomatic is the wide range of perceptual evidence that was accepted as confirming the existence of group anxiety. All are aspects of the syndrome evident in this type of situation.

As we have seen, the initiating stimulus complex is instability or ambiguity in one's environment. It may be a function of one's inability to identify objects or events, of inability to determine their meaning or cause, or to predict the outcome, or of inability to organize or plan the course of action to be taken. All of these imply the absence of relevant past experience to support an organized perception. Thus, among the subjects in these experiments in the high-anxiety categories, there were some who had formulated clear hypotheses concerning what was happening; but who had not evolved plans for action. For them the ambiguity was not so much in the meaning of events in the immediate situation as in the ambiguity of outcome or of action to be taken (15) as was shown by one subject who, having formulated a firm hypothesis that the events meant war, was in tears over the uncertain fate of her fiancé who was a member of the Reserve Corps. [See, also, Janis (17).]

The importance of having patterns of behavior available to serve as a stabilizing influence is illustrated by numerous examples. The emphasis upon training for emergencies, fire drill in schools, and aboard ships is in recognition of this principle. It has been reported that panic among troops subject to strafing by enemy aircraft during World War II was effectively reduced by instructing them to fire at the attacking planes with their rifles or sidearms—in spite of the fact that the probabilities were not large that such fire would be effective. It is a question as to whether this anxiety reduction was due to the action itself or was a function of the cognitive organization provided by knowledge of *action to be taken* in the situation. Troops in panic flight are taking action; yet the anxiety, according to anecdotal accounts, is not reduced—rather it appears to be enhanced.

It would appear that any stability in the present environment and one's place in it, whether it be narrow to include only a limited area of activity or broad to include a clear perception of the total situation, aids in reducing anxiety. Children evacuated from London during the war, for a time at the cost of breaking up families, clung desperately to their dolls or teddybears as the one structured element in their otherwise strange environments (7).

No doubt the emphasis upon training for meeting emergencies is correct, but possibly for reasons other than to develop behavior patterns *per se*. The

training may provide some stability in situations that otherwise would be perceptually or cognitively unstructured.

12. *The Problem-Solving Process*

Recounting the course of events in these experiments, the first response characteristically made was curiosity. The stimulus was the newsboys calling "Extra." As new elements were added to the situation, curiosity changed to anxiety. We can presume that both the curiosity and anxiety initiate the exploratory or problem-solving activity which for the most part took the form of covert hypothesis formation. Finally, both tension and activity ceased when the problem was solved; that is, when recognition occurred, which for many was delayed until they were told that it had been an experiment.

13. *The Issue of Age and the Response to the Unfamiliar*

The first question that arises concerning this sequence of events is the curiosity-anxiety response pattern. Its origin could be, as Hebb (12) has proposed, neither innate nor a function of learning; but nevertheless dependent upon both (6). On the other hand, although young animals do not display the signs of anxiety to those unfamiliar stimuli that alert an adult, their failure to respond may reflect the fact that by virtue of their limited experience a multitude of stimuli are novel, with the result that they are in a constant state of generalized adaptation (13, 14). This is a second possibility. Another, previously mentioned, concerns hypothesis specificity as a result of previous one-trial experiences and less rigorous criteria for recognition in the young organism. And finally a fourth relates to the fact that adult animals will respond with fear to an unfamiliar object if it is presented against the background of a familiar environment; but will be unresponsive to the same object in a less familiar setting, the situation which initially exists for the very young animal (14). This phenomenon has also been evident in these studies. In any case, with age the environment of the young becomes familiar and stabilized, so that novel additions to their surroundings are responded to with the arousal of the biologically important alerting mechanism of curiosity anxiety with exploratory behavior as a concomitant directed at identification.

The anxiety evident in isolation experiments discussed by Haggard (10) and the common fear of the dark have interesting parallels to the phenomena described here. Absent in isolation situations is stimulation that has connotations of threat. This, as we have seen, was not a necessary condition for the arousal of anxiety in these experiments. The reduction in sensory input during

isolation appears to allow the development of unorganized associative processes accompanied by anxiety, an observation that parallels the evidence described here that confusion of hypotheses is significantly associated with anxiety. Parallel also is the phenomenon of adaptation that occurs to both isolation and the unfamiliar.

14. *The Issue of Reinforcement*

The second question in this sequence concerns hypothesis formation, the covert exploratory behavior. These represented learned solutions to a problem, albeit for the most part erroneous in these situations. Here, as noted, exploration was initiated at the time of failure to identify events or their meanings. Referring to the underlying neurological activity it may be, as Hebb has proposed, a function of disruptions of phase sequences. This effort at recognition (to achieve closure) we infer is unlearned; and the tension reduction (the elimination of the nagging urge to identify) instead of serving as reinforcement for successful associations (that is, for learning) is simply another aspect at the physiological level of this closure. In other words, a stimulus that registers arouses associative processes, intrinsic parts of the perceptual process of recognition. Failure of these associative processes enhances these activities that we have identified as curiosity or anxiety, the "biogenic drives to know" (23).

These problem-solving activities are primitive and biologically important for survival. They resemble in one respect some of the problem-solving activities of the laboratory involving ambiguous stimuli: i.e., both require the same kinds of exploratory or hypothesis-forming behavior. They differ in that the tension compelling the effort may be in part or entirely of social origin in the socially imposed task, extrinsic to the stimulus; whereas in the primitive type of problem situation with which these studies were concerned the tension is a function of the ambiguity itself, the failure at recognition.

As the foregoing implies, and for the types of situations discussed, the tension reduction specifically associated with recognition or closure does not affect the learning *per se*. The subjects in these experiments later recalled both their discarded erroneous nontension reducing hypotheses as well as their few correct ones. In other words, although these problem-solving behaviors, overt or in the form of hypothesis formation, were a function of tensions, the learning that occurred was not associated with its reduction. It is the distinction between performance and learning where the latter is conceived as cognitive and not necessarily a function of reinforcement (27).

It is difficult to overestimate the importance of this kind of behavior (1). As

noted, we tend to identify its milder forms as reflecting curiosity and the associated tensions as exciting. It can be pleasurable, in other words, as well as disturbing (12, 28). The number of situations that depend at least in part for their interest and appeal upon the element of uncertainty is almost unlimited. The search for the novel, the exciting, the situations of uncertain outcome (although often combined with other sources of satisfaction) is the basis of most of our recreation. Games, travel, gambling, and in literature, the detective story and the novel derive their stimulus value in part or largely from this source. Advertisers early achieved 100 per cent uncertainty with the expression "Even your best friend won't tell you" and they have gone on from there. The searching out of situations that arouse these pleasant forms of mild excitement would appear to be a major preoccupation of many, and the number of industries to provide the varieties of entertainment that have this common characteristic is growing.

D. SUMMARY

This paper describes studies with human subjects which represent extensions of earlier work on the emotional behavior of animals in response to the strange or unfamiliar. The methods employed were in principle similar. Subjects in their otherwise normal environments were exposed to sounds which were ambiguous in terms of their instrumental origins or of their meaning and implications. As with animals, ambiguities in the environment have uniformly aroused responses in humans that lie along what appears to be a curiosity-anxiety continuum.

In one series of studies intermittent sounds of fire equipment and aircraft were presented to captive audiences over 25-minute periods. The data obtained were derived from moving-picture and observer records made during the experiment and from interviews and questionnaires following. The common interpretations made in these instances were war, threat of war, fire, and accident. Anxiety was typical, a response which was correlated with a number of other measures: with perception (projection) of anxiety in the group as a whole ($p < .01$), with suggestibility ($p < .10$), with "scapegoating" ($p < .05$), and with failure in retention ($p < .01$).

A covert pattern of responses that always occurs on exposure to ambiguous events is the formulation of hypotheses in attempts at recognition. These are rapid associative processes often involving remote and seemingly irrelevant experiences. Anxiety was correlated with their degree of confusion ($p < .01$), but *not* necessarily with the seriousness of the inferred threat. In the experiments cited, different methods were employed in efforts to relieve anxiety:

e.g., prayer and humor. Mild anxiety was aroused in some subjects by sounds of fire trucks and aircraft for several days after the experiments, responses which were related to the amount of anxiety aroused during the experiment ($p < .05$).

The learning associated with response to the strange is discussed from the point of view of traditional drive-reduction reinforcement-learning theory.

Methodologically this approach, in using ambiguous stimuli as the independent variable, might appear as a distant cousin of the projective test. This is not the case. It differs in the first place in that the subject's focus of effort is upon resolving the ambiguities through establishing stimulus identity. Although both involve associative processes, in these studies the associations are stimulus controlled and relate primarily to the subject's objective, environmental past, rather than to underlying motivational or subjective components. It also differs from the projective test in that the ambiguous stimulus is in fact the independent variable, and the response itself is the focus of interest. In the projective test on the other hand, it is the reverse: The independent variable is the unknown, to be inferred from the response; and the ambiguous stimulus patterns (e.g., the TAT card), a pseudo-independent variable, serves primarily as the medium or the occasion for the indirect expression of underlying motivational dispositions.

REFERENCES

1. BAKER, G. W., & CHAPMAN, D. W., Eds. *Man and Society in Disaster*. New York: Basic Bks., 1962.
2. BRUNER, J. S. Personality dynamics and the process of perceiving. In R. R. Blake & G. V. Ramsey (Eds.), *Perceiving—An Approach to Personality*. New York: Ronald, 1951.
3. CALDWELL, J. M., RANSON, S. W., & SACHS, J. G. Group panic and other mass disruptive reactions. *U.S. Armed Forces Med. J.*, 1951, 2, 541-567.
4. CANTRIL, H. *The Invasion from Mars*. Princeton, N.J.: Princeton Univ. Press, 1940.
5. ———. *The Why of Man's Experience*. Princeton, N.J.: Princeton Univ. Press, 1940.
6. DENNIS, W. Congenital cataract and unlearned behavior. *J. Genet. Psychol.*, 1934, 44, 340-350.
7. FREUD, A. *War and Children*. New York: Ernst Willard, 1943.
8. GRINKER, R., & SPEIGEL, J. P. *Men Under Stress*. Philadelphia: Blakiston, 1945.
9. GUETZKOW, H. Joining field and laboratory work in disaster research. In G. W. Baker & D. W. Chapman (Eds.), *Man and Society in Disaster*. New York: Basic Bks., 1962.
10. HAGGARD, E. A. Isolation and personality change. In P. Worchel & D. Byrne (Eds.), *Personality Change*. New York: Wiley, 1964.
11. HEBB, D. O. Emotion in men and animals: An analysis of the intuitive process. *Psychol. Rev.*, 1946, 53, 88-106.

12. ———. *Organization of Behavior*. New York: Wiley, 1949.
13. HUDSON, B. B. Extinction of the avoidance behavior of rats to strange objects. *Psychol. Bull.*, 1940, **37**, 591. (Abstract)
14. ———. One-trial learning on the domestic rat. *Genet. Psychol. Monog.*, 1950, **41**, 99-145.
15. ———. Anxiety in response to the unfamiliar. In D. W. Chapman (Ed.), *Human Behavior in Disaster: A New Field of Social Research*. *J. Soc. Issues*, 1954, **10**, 3.
16. JANIS, I. L. *Air War and Emotional Stress*. New York: McGraw-Hill, 1951.
17. ———. Psychological effects of warnings. In G. W. Baker & D. W. Chapman (Eds.), *Man and Society in Disaster*, New York: Basic Bks, 1962.
18. JONES, H. E., & JONES, M. C. A study of fear. *Childhood Educ.*, 1928, **5**, 136-143.
19. KRECHEVSKY, I. "Hypothesis" in rats. *Psychol. Rev.*, 1932, **39**, 516-532.
20. MEERLOO, J. A. M. *Patterns of Panic*. New York: Internat. Univs. Press, 1950.
21. MOWRER, O. H. *Learning Theory and Behavior*. New York: Wiley, 1960.
22. MURRAY, H. A. *Explorations in Personality*. New York: Oxford Univ. Press, 1938.
23. NISSEN, H. W. The nature of drive as innate determinant of behavioral organization. In M. R. Jones (Ed.), *Current Theory and Research in Motivation*. Lincoln, Nebr.: Univ. Nebraska Press, 1954.
24. RANSON, S. W. The normal battle reaction: Its relation to the pathological battle reaction. *Bull. U.S. Army Med. Dept.* (suppl.), November, 1939. Pp. 3-11.
25. SULLIVAN, H. S. *Conception of Modern Psychiatry*. Washington, D.C.: William Alanson White Psychiatric Foundation, 1947.
26. THORNDIKE, E. L. *Educational Psychology*. New York: Lemcke & Buechner, 1903.
27. TOLMAN, E. C. *Purposive Behavior in Animals and Men*. New York: Century, 1932.
28. WOODWORTH, R. S. *Psychology*. New York: Holt, 1921.

Department of Psychology
Rice University
Houston 1, Texas

SOME CORRELATES OF THE DON'T-KNOW RESPONSE IN OPINION QUESTIONNAIRES*¹

Center for Research on Social Behavior, University of Delaware

ROBERT C. ZILLER AND BARBARA H. LONG

A. INTRODUCTION

The opinion-rendering process may be viewed as a special case of decision making; a judgment is required under conditions of information limitations and precluding information search. Given these information restrictions, at least two components of the opinion process are suggested: the extent to which the respondent weighs the adequacy of his existing information resources as reflected in (a) the decision whether or not to withhold judgment or (b) the degree to which the respondent qualifies the judgment rendered. Operationally, the decision maker may or may not (a) choose the don't-know alternative in view of the limitations of information at his immediate disposal or, (b) if a judgment is made, the respondent may or may not prefer to employ a six-alternative opinion scale to a three-alternative scale (8). This paper is concerned with the first of these components.

In the area of decision making, a process comparable to this first component of the opinion process has recently undergone extensive analysis under the subject heading of information search (5, 6). The usual design permits the subject to gather additional information, usually at a given cost, before submitting a final decision. In the opinion process, however, it is usually assumed that the judgment must be submitted on the basis of existing information resources. Moreover, the possibility of the indefinite deferment of judgment is rarely an alternative in decision making situations.

A number of studies have been reported (2, 3, 4, 9, 11) which are relative to withholding judgment. Usually, however, the measures of withholding judgment involve attitude items (2, 9) or items from a personality questionnaire (4); and even though an I-don't-know alternative or a "?" alternative is included, information limitations are not a central consideration. Responses other than those indicating the withholding of judgment are implicitly requested.

* Received in the Editorial Office, Provincetown, Massachusetts, on April 14, 1964. Copyright, 1965, by The Journal Press.

¹ These studies were supported by the Air Force Office of Scientific Research Grant No. AFOSR 62-95. We are indebted to Drs. Berj A. Harootunian and Howard E. Lamb who assisted us in several phases of the study.

A measure of the tendency to withhold judgment under conditions of severe information limitations was reported by Ziller *et al.* (14). The experimental design involved a comparison of the don't-know responses of professional clinical psychologists and trainees in clinical psychology, who were presented with a psychiatric-observation form containing 180 items requiring descriptions and interpretations of the behavior of a character portrayed in a 27-minute movie skit. The total number of don't-know responses comprised the measure of the tendency to withhold judgment.

The measure of acknowledgement of knowledge inadequacy used in this study employed a response-set technique under conditions of information paucity to provide a disguised but a direct measure of the first proposed component of the opinion process. Of course, a measure with more general application requires more general content.

Withholding judgment under conditions of information inadequacy was assumed to be a component of dogmatism. However, because of the myriad of meanings associated with the term "dogmatism," this invitingly convenient concept was avoided in preference for the operational definition.

The two studies reported here were designed to analyze the relationship between Rokeach's Dogmatism Scale to withholding judgment and the relationship between withholding judgment and the respondent's age, sex, education, and intelligence.

B. STUDY I

In the previously described study by Ziller *et al.* (14), it was found that professionals as opposed to trainees in clinical psychology used the don't-know response less frequently in rendering professional judgments under conditions of extremely limited information and when a comparison was implied between the performance of professionals and trainees. It was assumed that under conditions permitting invidious comparisons between the decision-making behavior of professionals and trainees, the decision to defer judgment because of insufficient information was viewed by professionals as being incongruent with their perceived professional role. Furthermore, it was hypothesized that in a "highly subjective decision-making situation involving a status hierarchy, the high-status members are inclined to be more opinionated under conditions where the status relationship is salient." Opinionated or dogmatic behavior was interpreted as a mechanism for maintaining status or avoiding threats to status. The status of persons making the don't-know response was presumed to be more vulnerable than the persons making the I-know response, since

the don't-know response may suggest one or more of a variety of personal shortcomings. In this first study, the generalizability of the earlier findings to the conditions of age as a status variable was tested.

1. *Procedure*

The subjects were 120 male and female students attending the summer session at the School of Education at the University of Delaware. Both graduate and undergraduate students were included in the sample of subjects whose ages ranged from 19 to 59.

The opinion questionnaire was adapted from the instrument used in a study by Ziller and Mangano (13). The original 16 items were retained. Examples of the items are "There is life on other planets," "The world will always have wars," "Modern art will, in time, take its place beside classical art." The total number of don't-know responses provided the measure of withholding opinion under conditions of information inadequacy.

2. *Results*

The reliability of the withholding-judgment scale was estimated by means of a split-half technique. The resulting correlation coefficient corrected for length was 0.58 ($N = 120$).

The withholding-judgment scores were analyzed according to three age categories (under 30; 30 through 39; and 40 and above) and according to the sex of the respondent. These age groupings were designed to focus upon the middle age category, 30 to 39. Furthermore, earlier studies had revealed that this categorization system provided approximately equal numbers of subjects in each category with regard to this particular population (12). The requirements of proportionate number of subjects in each of the six cells of the age-by-sex analysis of variance dictated the elimination of 30 subjects through random procedures. The subjects retained represented the maximum number of subjects available in the sample under the statistical requirement of proportional number of subjects per cell. All subjects were either teachers or prospective teachers.

The results with regard to the main effects of age and the age-by-sex interaction were statistically significant ($p < .02$ —see Tables 1 and 2). The youngest group (those under 30) elected the don't-know response most frequently; whereas the middle age group (30 to 39) chose this response least frequently. The age-by-sex interaction was most pronounced between the youngest and oldest age categories. In the youngest group, the females

TABLE 1
MEANS OF THE FREQUENCY OF DON'T-KNOW RESPONSES
ACCORDING TO SEX AND AGE

Sex	Under 30		30-39		Above 39		Total	
	N	Mean	N	Mean	N	Mean	N	Mean
Male	12	2.58	12	1.50	12	3.42	36	2.50
Female	18	4.00	18	2.44	18	2.22	54	2.88
Total	30	3.43	30	2.06	30	2.70	90	2.72

TABLE 2
ANALYSIS OF VARIANCE OF THE FREQUENCY OF DON'T-KNOW RESPONSES

Source of variance	df	Mean square	F	p
Age	2	14.00	4.22	.02
Sex	1	4.00	1.20	—
Interaction	2	13.50	4.07	.02
Error	84	3.32		

submitted the greatest number of don't-know responses; in the oldest group, the males submitted the greatest number of don't-know responses.

3. Discussion

The results demonstrated that the middle age group selected the least number of don't-know responses more often than the other age groups. To the extent that the 30-to-39 age group of teachers may be assumed to be the most concerned about status or power or potential power in their education organizations, the results may be interpreted as supporting the initial hypothesis relating status defense and dogmatism.

The interaction effects between age and sex with regard to the don't-know response was paradoxical. In the age group under 30, the males made fewer don't-know responses than females; in the age group above 39, the females made fewer don't-know responses than males. Insufficient information concerning the differences in the samples of subjects in the two age groups counsels the authors to withhold judgment.

C. STUDY II

The results of the first study were interpreted in terms of the hypothesis relating status and dogmatism; but the results were scarcely conclusive. In the second study, several shortcomings of the initial study were corrected; the sample of subjects and the length and content range of the opinion questionnaire were increased. In addition, the relationships involving withholding judgment were analyzed with regard to Rokeach's Dogmatism Scale

and intelligence. Finally, the variable of educational level (freshman, sophomore, junior, senior) was substituted for the somewhat indistinct status variable of age. The sex categorization of subjects was retained in the analysis.

In a recent study (1) it was hypothesized that there is a "progressive decline in dogmatism with advancing age from Grade 8 to 12." The hypothesis was derived from Loevinger's theory of ego development (7). She contended that the adolescent learns to "assume distance from oneself." Thus, in early adolescence, the self-appraisal tends to be stereotyped, behavior which is also characteristic of dogmatism. Using an abbreviated form of Rokeach's Dogmatism Scale, the hypothesis was supported.

1. *Instrument Revision*

Analysis of the withholding-judgment instrument used in Study I revealed that future-oriented items elicited the greatest number of don't-know responses. Thus, in increasing the length of the instrument for purposes of improved reliability and validity, the new items were largely future oriented. Moreover, in an effort to avoid a possible subject-matter bias, we added a number of items from a wide variety of subjects including the following: (a) By 1980 the world's food supply will come largely from the sea; (b) The world economy would eventually be improved if all tariffs were eliminated; (c) Children who move frequently are not as well adjusted as those who grow up in one place; (d) By 1980, 90 per cent of the American population will be living in cities of 50,000 or more.

A total of 38 items comprised the general opinion questionnaire used in Study II. The subjects were asked to circle one of the three alternatives with regard to each statement: (a) Agree, (b) Disagree, (c) Don't know. Again, the total number of don't-know responses comprised the withholding-judgment score.

2. *Validation as a Measure of Dogmatism*

Heretofore it has been assumed that withholding judgment under the conditions of the questionnaire already described is related to dogmatism. In order to test this assumption, withholding-judgment scores were related to scores on Rokeach's Dogmatism Scale. With regard to a sample of college women ($N = 72$), $r = -.32$ ($p < .01$); with regard to a sample of summer-school students, largely teachers ($N = 52$), $r = -.27$ ($p < .05$); and with regard to a sample of elementary-school teachers ($N = 44$), $r = -.47$ ($p < .01$).

Although these correlation coefficients are low, they are consistent and statistically significant. In comparison with the Rokeach Scale, which derives from an inchoate theory of personality, the withholding-judgment scale has the advantage of being operationally defined, objective, and emanating from a low-visibility response-set approach to personality measurement.

3. *The Relationship of Withholding Judgment to Intelligence and Academic Status*

In the second study proper, the subjects were 589 regular, undergraduate, male and female students from the departments of education and history at the University of Delaware. Included in the sample were 165 freshmen, 173 sophomores, 147 juniors, and 104 seniors.

4. *Results*

In exploring the relationship of the two opinion-process measures with intelligence, the correlation coefficient of the verbal subtest of the SCAT and grade-point average with the withholding-judgment scores of the freshmen women sample ($N = 134$) were -0.14 ($p < .10$) and -0.24 ($p < .01$) respectively. Thus, more capable students tended to select the don't-know alternative less frequently; but the results were more highly related to the more visible indicator of ability: grade-point average.

With reference to the analyses involving educational level and the sex of the respondent, simple analyses of variance with respect to educational level were calculated separately for males and females since a small and disproportionate number of males were included in the sample. The results are reported in Table 3.

The results with regard to males and females were very similar but were statistically significant only with respect to the females ($F = 4.08$; $df = 3$,

TABLE 3
MEANS OF THE FREQUENCY OF DON'T-KNOW RESPONSES
ACCORDING TO EDUCATIONAL LEVEL AND SEX

Educational level	Males*		Females**		Total***	
	N	Mean	N	Mean	N	Mean
Freshman	31	10.9	134	10.1	165	10.3
Sophomore	46	9.2	127	9.7	173	9.6
Junior	34	8.0	113	8.1	147	8.1
Senior	37	10.0	67	9.5	104	9.6

* $F = 2.05$, $p < .10$.

** $F = 4.08$, $p < .01$.

*** $F = 5.78$, $p < .001$.

440; $p < .01$). Pooling the data regarding both males and females, the results were also statistically significant ($p < .001$).

The results describe a decreasing frequency of the use of the don't-know response through the junior year in college. In the senior year, however, the subjects were found to use the don't-know response more frequently again. Overall, however, the statistical results were largely attributed to the difference in the responses of juniors who differed significantly from members of all the other classes.

D. DISCUSSION

The significant and consistently negative relationships between a withholding-opinion index and Rokeach's Dogmatism Scale support the hypothesis that avoidance of the don't-know response under conditions of information inadequacy is an indication of dogmatism.

The results of the studies involving professionals and trainees in clinical psychology, teachers in various age groups, and college students varying in academic status indicate that persons possessing relatively high status under conditions of instability or threat to the group structure tend to respond to an opinion questionnaire in accordance with perceived status and group-position expectations concerning their knowledge. More succinctly, it is proposed that dogmatism is a status-defense mechanism. Thus, clinical psychologists, when their responses were to be compared with trainees in clinical psychology, used the don't-know response less frequently under conditions when "don't know" was the most appropriate response. Teachers in the 30-to-39 age group as compared with younger and older teachers tended to eschew the don't-know response under conditions when they presumably possessed insufficient information upon which to render a reliable opinion. Then, too, college students at the junior level tended to use the don't-know response less frequently than members of the other classes under conditions similar to those under which the preceding results were observed. Finally, the results relating two academic measures of ability with the withholding-judgment index suggest, albeit highly speculatively, that the highest negative correlation ($r = -.24$, $p < .01$) between mental ability and dogmatism was found with regard to the most patent index of academic ability (assuming that grade-point average is more widely known and understood throughout high school and in later academic years than SCAT achievement scores). Here, too, then, it is proposed that the tenuous status with regard to school achievement is defended by conforming to the expectations of a superior student; that is,

the expectation that a superior student is knowledgeable and therefore need not have recourse to the don't-know response.

The marked rise in the don't-know responses when comparing juniors and seniors is especially remarkable in view of the attenuation effects of the greater intelligence of seniors through selective screening processes. (It will be recalled that a low negative correlation was found between intelligence and frequency of the don't-know response.) It is proposed that the diminishing frequency of the don't-know response from the freshman through the junior year reflects the student's perception of status within the college community. The increase in don't-know responses in the senior year is interpreted as reflecting the senior student's increasing awareness of his impending transition from the transient college community to the society in which he must find a relatively permanent position. Thus, it is proposed that the senior student responds to the opinion questionnaire in terms of his newcomer role with regard to the host society rather than in terms of his status within the transient donor college community. Although the senior student has high status by virtue of his education level within the college community, within the working community of which he is an imminent newcomer, a college education is often only a basic requirement of admission.

In conclusion, the results lend support to the component approach to the study of the opinion process and indicate caution in the interpretation of questionnaire data in which a high percentage of don't-know responses is elicited. It is proposed that there exists a positive correlation between the percentage of don't-know responses and the probability of introducing a response set along with the inherent personality correlates.

E. SUMMARY

Two experiments were reported investigating some of the correlates of the don't-know response in a general opinion questionnaire. The don't-know response or withholding judgment was proposed as basic component of the decision-making process involved in questionnaire-response behavior. The first study involved male and female teachers in three age groups: under 30, 30 to 39, and 40 and above. The second study involved freshman, sophomore, junior, and senior college students of both sexes. It was found that the middle age group and students in their junior year used the don't-know response least.

Interpreting the avoidance of the don't-know response as an indicator of dogmatism, it was proposed that dogmatism is a status-defense mechanism. In addition, the relationships are reported of withholding judgment with

Rokeach's Dogmatism Scale and measures of intelligence. Implications for the interpretation of questionnaire data are discussed.

REFERENCES

1. ANDERSON, C. C. A developmental study of dogmatism during adolescence with reference to sex differences. *J. Abn. & Soc. Psychol.*, 1962, **65**, 132-135.
2. DUNETTE, M. D., UPHOFF, W. H., & AYLWARD, M. The effect of lack of information on the undecided response in attitude surveys. *J. Appl. Psychol.*, 1956, **40**, 150-153.
3. EDWARDS, A. L., & DIERS, C. J. Social desirability and conflict. *J. Soc. Psychol.*, 1962, **58**, 349-356.
4. FONDA, C. P. The nature and meaning of the Rorschach white space response. *J. Abn. & Soc. Psychol.*, 1951, **46**, 367-377.
5. IRWIN, F. W., SMITH, W. A., & MAYFIELD, J. F. Tests of two theories of decision in an expanded judgment situation. *J. Exper. Psychol.*, 1956, **51**, 261-268.
6. LANZETTA, J. T., & KANAREFF, V. T. Information cost, amount of payoff, and level of aspiration as determinants of information seeking in decision making. *Behav. Sci.*, 1962, **7**, 459-473.
7. LOEVINGER, J. A theory of test response. In *Invitational Conference on Testing Problems*. Princeton, New Jersey: Educational Testing Service, 1958.
8. LONG, B. H., & ZILLER, R. C. Qualification of judgment: A learned cognitive style. *Percept. & Mot. Skills*, 1964, **18**, 235-238.
9. MITZELL, H. E., ROBINOWITZ, W., & OSTREICHER, L. M. The effects of response sets on the validity of Minnesota Teacher Attitude Inventory. *Educ. & Psychol. Meas.*, 1956, **16**, 501-515.
10. ROKEACH, M. *The Open and Closed Mind*. New York: Basic Bks., 1960.
11. ROSENBERG, N., IZARD, C. E., & HOLLANDER, E. P. Middle category response: Reliability and relationship to personality and intelligence variables. *Educ. & Psychol. Meas.*, 1955, **15**, 281-290.
12. ZILLER, R. C., & EXLINE, R. V. Some consequences of age heterogeneity in decision-making groups. *Sociometry*, 1958, **21**, 198-211.
13. ZILLER, R. C., & MANGANO, D. Cognitive style and interpersonal perception. Unpublished paper, University of Delaware, Newark, 1962.
14. ZILLER, R. C., SHEAR, H., & DECENCIO, D. Dogmatism—A professional response set. *J. Clin. Psychol.*, 1964, **20**, 299-303.

Center for Research on Social Behavior
University of Delaware
Newark, Delaware

PERCEPTUAL ACCENTUATION AS A PROBLEM IN THE PSYCHOLOGY OF JUDGMENT* 1, 2

Department of Psychology, University of Miami

RAYMOND E. HARTLEY

A. INTRODUCTION

It is likely that the most widely believed hypothesis for the explanation of perceptual accentuation is the one that attributes the overestimation on the part of the subject to an effect upon his phenomenal experience; that is, he overestimates because the object *looks* larger. Yet if we look at the current evidence we find not a single study of accentuation in which the experimenter has attempted to get reasonably direct evidence that the effect is visual. But surely, considering the very considerable gap between the subject's presumed phenomenal experience and his response, there must be ample scope for more plausible (though more pedestrian) explanations.

For hints toward such alternate explanations we would do well to examine the experimental paradigm of the accentuation studies. It is easy to state. The subject is required to judge some attribute, call it "A," of an object. Over a series of judgments the experimenter systematically varies some other attribute of the object, call it "B." Accentuation is shown when the variation in B has affected the judgments of A in comparison to what they would have been if B had been held constant. In current experimentation B, which we might call the "disturbing variable," is usually the value or the interest of the object. This restriction is not necessary—B might as well be any variable the experimenter may find of interest.

From this standpoint it is plain that we are dealing with a problem in psychophysics. These are studies that are busy examining the effect of a disturber variable upon the judgments of another variable. The general strategy of such investigation should address itself to the following points:

1. Do disturber variables affect judgments?

* Accepted by Raymond B. Cattell of the Editorial Board, and received in the Editorial Office, Provincetown, Massachusetts, on April 14, 1964. Copyright, 1965, by The Journal Press.

¹ This study was made possible by a grant from the Johnson Fund of the American Philosophical Society.

² I wish to thank Mr. William Wimsatt and Mr. Martin Miller for their help in the design of the experiment and Mr. Robert Hunter and my wife, Leah, for their help in the carrying out of the experiment and the consequent computational work.

2. If so, are some types of disturber variables more effective than others? Perhaps some types are not effective.

3. Having found an effective type of disturber, what is the precise statement, or law, of its effect?

Before the first perceptual-accentuation study made its appearance, Point 1 could have been answered in the affirmative. It has long been known that a disturber variable affects the judgment of some other variable. When judging weights by the method of constant stimuli it was customary to use several standard weights because the standard weight would warm toward body temperature through repeated handling, and it was known that difference of temperature would affect the judgment of weight. The older psychophysics was perfectly persuaded that any extraneous property that was allowed to vary would affect the judgments and routinely avoided these effects by requiring that the objects to be judged be identical in all respects save the one property that was of interest.

The particular contribution of the perceptual-accentuation studies has been to introduce as disturber variables value or interest instead of the sensory variables the effects of which the older psychophysics was so well aware. In addition, this alteration implies a wish to study the disturbing effect; whereas the early psychophysics wished to eliminate such disturbances, not study them. This changed emphasis brings Points 2 and 3 to the fore. Since value, as a possible source of disturbance, seems to be a very different matter than a sensory variable, such as temperature or color or size, it immediately becomes a question for investigation as to whether both types of disturbers disturb. We can assume that the older psychophysics showed that sensory variables serve as disturbers, and I believe that the perceptual-accentuation studies have shown that social variables also serve as disturbers. If so, the weight of interest should now turn to Points 2 and 3.

To approach these questions, I believe that the paradigm for an accentuation study should be changed from its present qualitative form to a quantitative form by requiring that the judgments be made in such a way as to be subject to psychophysical scaling. We could then get the scale values for the A variable when the disturber variable was absent, get them again (call them A') when the disturber variable was present, and scale the disturber variable itself. The precise answer to Points 2 and 3 would be at hand when the best fitting form of the function

$$A'_j = f(A_j, B_j) \quad [1]$$

was known. Here " A " and " A' " are as specified above, " B " is the value of the disturber variable, and " j " is the designation of the object being judged. To express the function precisely would be the answer to 3, and the answer to 2

would follow. If, for instance, [1] had a different form when B was a sensory variable than it did when B was a social variable we would know at once that 2 was incorrectly expressed. Social variables would be seen to be operating differently than sensory variables. If the form of [1] was the same in both cases then the appropriate constants of [1] could be used as measures of relative efficiency.

I argue, then, that the current situation in the field of perceptual accentuation necessitates that we turn our attention to the discovery of the form or forms of [1]. Let us drop our efforts to show that social variables do disturb and try instead to determine the manner in which the disturbance is accomplished. The present study is meant to exemplify this intent. The variable "A" to be judged is the area of geometric figures; "A" is to be the area when judged in the presence of a disturber "B"; "B" is to be judged monetary value associated with the figures; and the primary object is to specify the form of [1].

B. PROCEDURE

The subjects used were 124 volunteers from introductory-psychology classes at the University of Miami. They were assigned to one of four groups, two experimental and two control, in rotation as they appeared for testing.

It was decided that a convenient "A" variable for judging would be the area of simple geometric silhouette figures. As a convenient source these were selected from Birkhoff's *Aesthetic Measure* (1). While perhaps of no importance as a control, the figures selected were of the same aesthetic measure. They are shown in Figure 1.

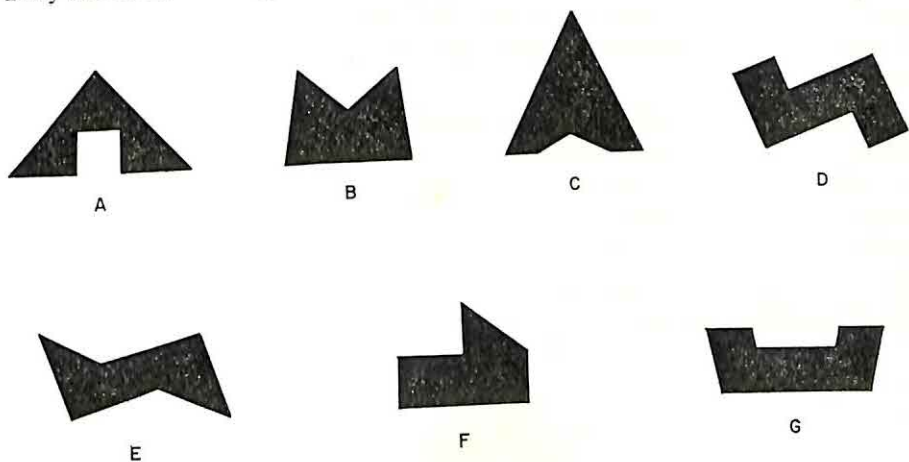


FIGURE 1
THE SEVEN BIRKHOFF FIGURES WHICH WERE USED AS STIMULI

The judgment of area was to be done by the method of pair comparisons; so the materials were prepared as follows. It was decided that the figures were to differ in area in accord with the following rank order: E, B, F, A, G, C, D. This is in order from small to large. The figures were drawn on separate sheets and with approximately equal increases of area each step up the rank order. The 21 possible pairings of these seven figures were then photographed for two-by-two slides.

1. *The Control Group*

The control group was required to judge, for each of the 21 slides, which of the figures was larger. The slides were projected by an SVE Schoolmaster projector on a screen 40 cm away. An Airequipt magazine was used with the projector to keep the slides in a fixed order for projection.

By means of a shutter all slides were exposed, throughout all parts of the experiment, for one second each. This standardized the exposure of a given stimulus to all subjects.

One-half ($N = 31$) of the control subjects were required, for each of the 21 presentations, to signify by saying "right" or "left" which of the figures they thought was the larger (the greater-than group). The other half of the control subjects signified by "right" or "left" which of the two figures they thought was the smaller (the less-than group).

This judgment of area for 21 pairs was the complete task for the control subjects. It was also the last task performed by the experimental subjects.

2. *The Experimental Group*

Before performing the last task of judging area the experimental subjects underwent a training procedure that was intended to cause them to regard these seven figures as of various monetary values. This procedure was set up as follows.

The figures were redrawn, all on separate sheets, and were all drawn to be the same area. An 84-presentation pair-comparison schedule was drawn up in accordance with the recommendations of Ross (4). This schedule was prepared for slide projection as before.

Each experimental subject was given 50 cents at the beginning and told that this was his to keep as well as was any further money that he might win in the course of the concept-formation experiment to follow. For the concept-formation experiment, monetary value had been assigned to each of the seven figures as follows. In a pair-comparison schedule of 84 presentations with seven objects, each object will appear on 24 of the presentations. The monetary value designated for a given object was a variable. Each figure was assigned

an average monetary value and a distribution of values over which its monetary value could range. This distribution of values was the same for all figures, so that in its 24 appearances a given figure had its average value assigned six times; a value two cents above its average, four times; a value two cents below its average, four times; a value five cents above, two times; a value five cents below, two times; a value nine cents above, two times; and a value nine cents below, two times; a value 13 cents above, once; and finally a value 13 cents below, once. The average values were set so: for A, 14 cents; for B, 16 cents; for C, 18 cents; for D, 20 cents; for E, 22 cents; for F, 24 cents; and for G, 26 cents. The rank ordering of the objects for area was set so as to correlate zero with this ordering of average monetary value. Finally, the foregoing distribution of values for a figure was assigned to the 24 appearances of the figure in a random fashion.

After being given his initial 50 cents, the subject was told that he would see a series of pairs of figures and that on each pair he was to guess which was the more valuable. If he guessed correctly he would be given the difference in their values; if he guessed incorrectly the difference in their values would be taken away. It was his job to learn what the values of the figures were, and obviously the faster he did the more he could win.

Although the subjects tended to look for complex rules that would enable them always to guess correctly, it is clear that the problem was so set up that all they could achieve was to form an impression of the rank order of average values. The forming of this impression was the aim of the "experiment," and after the first 10 subjects a standard procedure of continuing each subject through the series of 84 as many times as necessary for him to win three dollars was adopted. By the time a subject had won three dollars he seemed to have a good but not letter-perfect knowledge of the rank order of value.

After the subject reached the criterion of winning three dollars, he was told that the experiment was done satisfactorily. The projector was then returned to the beginning of the series of 84 and, without any exchange of money the subject was asked to signify (for the first 21 slides) which of the pairs was the more valuable. Having, hopefully, built up the concept of value in the subject by the preceding training the purpose was to make observations on the nature of each subject's concept of that value.

Finally, each subject judged areas—as had the control subjects. At this point the experimental subjects were divided into two groups as with the control subjects: the greater-than and the less-than groups. Thus the control group, divided into two, judged the area of the seven figures. The experimental group first experienced a concept-formation experiment designed to cause them

to attach monetary values to the figures. After which they judged these figures for area as did the control group.

C. ANALYSIS OF THE POOLED GROUPS

As the first step of analysis, the two control groups were pooled as were the two experimental groups. We thus begin with three tables of proportions.³ One table gives the proportion of judgments "greater than" for the monetary values of the seven stimuli, as given by the experimental group. One table gives the proportion of greater-than judgments for area, as given by the control group. And one table gives the proportion of greater-than judgments for area, as given by the experimental group. The less-than judgments given by half of each group were reversed for pooling.

These data were scaled by the method of pair comparisons under Case III of the Law of Comparative Judgment (3, Chapter 7). These scale values are presented in Table 1.

TABLE 1
SCALE VALUES FOR AREA AND MONETARY VALUES FOR THE POOLED EXPERIMENTAL AND CONTROL GROUPS

Object	Area Experimental	Area Control	Value Experimental
A	1.033	.459	0
B	1.062	.987	1.323
C	2.508	2.282	1.075
D	2.123	1.920	1.796
E	0	0	1.910
F	.775	.677	2.654
G	1.614	1.375	3.471
χ^2*	9.52	18.00	6.40

* At the .01 level and nine degrees of freedom, chi square = 21.666.

The chi-square entries in the table are the chi-square values resulting from the Mosteller consistency test (3, pp. 167-168). Since these do not reach the one per cent level of significance, which is recommended by Guilford, it appears that these scale values are consistent with the obtained proportions.

The scale values for area differ considerably between the experimental and control groups. This, of course, is exactly what we are interested in, the first problem being how to test this for significance? Since the two sets of scale values are derived from tables of proportions, what we need is a test that gives

³ These tables have been deposited as Document number 8417 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. A copy may be secured by citing the Document number and by remitting \$1.25 for photoprints, or \$1.25 for 35 mm microfilm. Make checks or money orders payable to Chief, Photoduplication Service, Library of Congress.

an overall test of significance between two tables of proportions. Fundamentally, the Mosteller test of consistency does this. This test tests the significance between a set of proportions obtained from the scale values and the obtained proportions.

For the case in which both sets of proportions are obtained proportions it appears that the degrees of freedom are:

$$\frac{k(k-1)}{2} = 1 \quad [2]$$

where k is the number of stimuli.

With seven stimuli the degrees of freedom as given by [2] are 20. The chi square between the two tables of proportions is 50.112. This, then, is significant beyond the .01 level.

Now, what will be a plausible form for [1] to describe an accentuation effect? The following seems to express the customary view concerning perceptual accentuation:

$$A'_j = eA_j + dB_j + k \quad [3]$$

Here the A' values are regarded as basically linear transformations of the A values plus a small increment from the judged monetary value B . The quantity dB_j , the accentuation term of [3], causes the values A' to be inflated over what they would have been, and this inflation is larger the larger B is. It was this type of relation that Bruner and Goodman (2), in the original coin-judging experiment, thought they had discovered when they plotted the amount of overestimation against the values of the coins and found, with the exception of the 50-cent piece, an increasing function of overestimation with value.

We need, now, a method of determining e , d , and k that gives a best fit of [3] to the whole data. Noticing that [3] is in the standard form of a multiple-regression equation we compute the correlations $r_{A'A}$, $r_{A'B}$, and r_{AB} over the seven stimuli as the sample and proceed in the usual fashion.⁴

The resulting equation, for our data is:

$$A'_j = 1.110 A_j + .031 B_j - .059 \quad [3a]$$

While [3a] is the best fit that an equation of the form of [3] can make to these data it may, nevertheless, not be a good fit. It appears that this matter can be settled as follows. If, for a given stimulus, we substitute in [3a] the obtained values for the A_j and B_j the equation yields a predicted or "theo-

⁴ This method was suggested by Quinn McNemar.

retical" value of A'_j . These seven "theoretical" A'_j values appear in the second column of Table 2 (the first column of Table 2 contains the obtained A'_j values for comparison). Now we take the "theoretical" scale values derived from [3a] and, behaving as if they were our obtained scale values, we work

TABLE 2
PREDICTED SCALE VALUES FOR A'_j DERIVED FROM EQUATIONS (3a), (4a), (4b)

Object	Original	A'_j values		
		(3a)	(4a)	(4b)
A	1.033	.450	1.029	.499
B	1.062	1.078	1.094	1.074
C	2.508	2.507	2.522	2.483
D	2.123	2.128	2.092	2.089
E	0	0	0	0
F	.775	.775	.752	.737
G	1.614	1.575	1.623	1.496
χ^2 *		36.61	8.89	34.90

* At the .01 level and nine degrees of freedom, chi square = 21.666.

the scaling procedure backwards to get a set of theoretical proportions implied by these scale values. This set of proportions is then tested against the actual set of proportions, by Mosteller's test. The chi square is given in Table 2 and is well beyond the .01 level. Clearly, then, these theoretical scale values do not agree with the empirical proportions and [3a] cannot be regarded as giving a satisfactory fit to the data.

Thus the customary concept that the amount of accentuation is proportional to the value of the object has failed in the context of this experiment. The question now is whether a more complex function of B can be found that will do the job.

Inspection of the data led to the impression that value was effective, as a disturber, as a function of its "extremeness" rather than the absolute value. Furthermore, it seemed that the effect was not merely proportional to the degree of "extremeness" but increased rapidly with increasing extremeness. This impression was implemented as follows.

The central tendency of the B s was computed. Here the value of Object A was quite deviant; so it was dropped and the central tendency was taken as the mean of the remaining six values. This mean was 2.038. The absolute deviate scores of each object were then figured and appear in Column 1 of Table 3 under the heading E^*_j . These scores were supposed to represent "extremeness" of value. The antilogarithms, to base 10, of these E^*_j values were then determined and appear in Table 3 under the heading Q_j . The

TABLE 3
CERTAIN FUNCTIONS OF THE B_j VALUES AS USED IN EQUATION (4A)

Object	E_j^*	Q_j
A	2.038	109.100
B	.715	5.188
C	.963	9.183
D	.242	1.746
E	.128	1.343
F	.616	4.130
G	1.433	27.100

intention of these Q_j values was to represent the supposed increasing effectiveness of the more extreme values. The law to be tested is, then,

$$A'_j = eA_j + dQ_j + k \quad [4]$$

Although the Q_j values are certainly not linearly related to the B_j values, still [4], itself, is linear and can be fitted to the data just as before. We get

$$A'_j = 1.088 A_j + .0049 Q_j - .005 \quad [4a]$$

The predicted scale values given by [4a] appear in Table 2 and, by the Mosteller test, are certainly a good fit to the data.

In view, however, of the small absolute size of d we must wonder if the accentuation term is necessary. Perhaps only the linear transformation portion is needed and the accentuation term adds nothing of importance. In other words, the abridged equation [4b] is all that matters.

$$A'_j = 1.088 A_j - .005 \quad [4b]$$

The "theoretical" A'_j values obtained from [4b] are in the last column of Table 2. The chi square of the Mosteller test is beyond the .01 level; hence [4b], by itself, is not sufficient and we can regard the accentuation term of [4a] as significant.

D. ANALYSIS OF THE GREATER-THAN AND LESS-THAN GROUPS

The experimental group was composed of subjects one-half of whom, on the area judgments, were asked to declare which of the two figures was the *greater*; and the other half were asked to judge which of the two figures was the *smaller*. The control group was divided similarly. Since the pooled groups show a significant difference, it becomes of interest to see if this procedure produced any differential effect.

The four tables of proportions were scaled by Case III.⁵ The scale values obtained are given in Table 4. The chi-square entries in Table 4 give the results of the Mosteller consistency test. Since none is significant, it appears that Case III scaling is adequate.

TABLE 4
SCALE VALUES FOR "GREATER THAN" EXPERIMENTAL AND CONTROL GROUPS AND
FOR "LESSER THAN" EXPERIMENTAL AND CONTROL GROUPS

Object	Greater		Less	
	Experimental	Control	Experimental	Control
A	.758	.234		
B	.859	.965	.890	.455
C	1.770	2.185	1.013	.781
D	1.635	2.124	2.504	1.945
E	0	0	2.270	1.666
F	.396	.594	0	0
G	1.288	1.304	.603	.476
χ^2 *	15.528	12.048	1.709	1.251
			7.150	16.84

* At the .01 level and nine degrees of freedom, chi square = 21.666.

The significance tests were managed as before, with the Mosteller test, using 20 degrees of freedom. For the less-than groups the chi square between the experimental and control groups was 66.348. With 20 degrees of freedom, significance at the .01 level demands a chi square of 37.566; hence, this is significant. The difference between the experimental and control groups of the greater-than group yields a chi square of 37.184 which is annoyingly marginal, but which we regard as close enough to be considered significant.

Due to the experience with the pooled groups it was supposed that [4] would fit the data. Since the judgments of monetary value given by the experimental group were entirely greater-than judgments there is no breakdown here, and the Q_j values to be used here are the same as before.

Fitting [4] to the present data as before we get for the less-than groups

$$A'_j = 1.315 A_j + .0027 Q_j - .012 \quad [5]$$

While this is the best fit that [4] can give to the data of the less-than groups, is it a good enough fit? The A'_j values predicted by substituting the obtained values in the right of [5] give a Mosteller consistency test against the less-than experimental proportions of 7.228. This is practically the same as given by the actually obtained scale values from these proportions. Thus it appears that [5] is a good fit to the data.

⁵ See Footnote 3, p. 154.

Is the accentuation in [5] of any value? After all, the value of .0027 seems small. Dropping the accentuation term we get

$$A'_j = 1.315 A_j - .012 \quad [5a]$$

The A'_j values predicted from this equation, when applied to the experimental group proportions, yield a Mosteller consistency test of 11.182. Now while this is higher than [5] yielded, these A'_j values from [5a] are still consistent with the actual proportions; hence we do not need to suppose that the accentuation term in [5] is significant.

For the greater-than groups we get

$$A'_j = .798 A_j + .0056 Q_j - .013 \quad [6]$$

The A'_j scale values predicted from [6] fit the greater-than experimental-group proportions virtually as well as do the scale values actually obtained from these proportions. The Mosteller consistency test applied to these predicted A'_j values gives a chi square of 15.656, which is much the same as yielded by the scale values actually obtained from these proportions.

When we drop the accentuation term from [6] we get

$$A'_j = .798 A_j - .013 \quad [6a]$$

This gives a consistency test against the experimental-groups proportions of chi square = 30.416. This is a very considerable change from [6] and, with nine degrees of freedom is well beyond the .01 level. It appears that for [6] to give a satisfactory fit to the experimental proportions the accentuation term is necessary.

E. DISCUSSION

Possibly, what happens is as follows. The subject is called upon to judge which of two objects is greater in some aspect, such as size; and the actual difference in size is slight. That is, the judgments the subject makes involve considerable doubt and difficulty on his part. [How *would* you design an accentuation experiment in which all the judgments required were so clear-cut as to cause no difficulty?] As the subject struggles to decide, he also has knowledge of or can see that one object is greater than the other in some other dimension. As he vacillates in his judgment, he perhaps swings momentarily toward one decision and since, to his knowledge, that object is *greater* (although in some other dimension) this makes his judgment *feel* right and he commits himself.

In more general terms, the argument is that when judgments are being made in conditions of considerable natural variability a contaminating variable will serve to bias the variability. It is, of course, now hypothesized that this biasing effect occurs in a way describable by [4] where e and k are the customary constants of a linear transformation; and d serves as an index of discrimination. The larger the value of d the poorer the subject discriminated A from Q . The presence of the Q 's in [4] rather than the original B s is, of course, of considerable interest. One should never forget that a high degree of relativity occurs in psychophysical judgment in that all the judgments are comparisons of one thing with another. Here the effect of the relativity upon the disturber variable is to establish a central tendency for the whole set of objects, with the disturber variable causing its difficulties in discrimination in terms of its deviation from this central tendency.

In the present study, for example, the difference between the greater-than and less-than groups would appear to offer other hypotheses considerable difficulty. If, for instance, the accentuation is caused by a genuine visual effect (that is, the object actually "looks larger" because it is more valuable and the subject is merely reporting what he sees), then I see no reason why the less-than judgments should give a different result from the greater-than judgments. If, however, accentuation is a matter of a bias being imposed on the subject's vacillations in response, the difference seems predictable. First we note that the relations "greater than" and "less than" are converses of each other. To know one is to know the other, but from a psychological point of view they are hardly equivalent. The judgment "greater than" is likely to be more natural to a subject than the judgment "less than." As a matter of fact, a number of subjects of the present study expressed such feelings. Very possibly, then, this happened in the less-than groups. The control group simply has some extra trouble in making the less-than judgment of area as compared to the difficulty it would have had if the judgment had been "greater than" but the increased difficulty does not give rise to any bias. The less-than experimental groups experience a conflict. They struggle to judge "less than" in area and are disturbed by their knowledge of which object is "greater than" in monetary value. But "greater than" is the opposite of "less than" and there will be conflict (not cooperation) between the two tendencies toward response. Since our assumption was that it was cooperation between the two tendencies that produced the bias we have no special reason here to expect any bias or, at least, to expect less bias. These guesses seem to be borne out in the present data. The accentuation effects seem to be mostly or entirely due to the greater-than groups. The d 's were:

.0027 less than
 .0049 pooled
 .0056 greater than

The "less than" was not significant, the pooled group was significant, and the d of the greater-than groups was significant. Apparently our observed accentuation effect was mainly if not entirely produced by the greater-than groups. It is of interest that the e 's are:

1.315 less than
 1.088 pooled
 .798 greater than

Since the unit of measure in pair-comparisons scaling is a standard deviation, these e 's may indicate that the less-than experimentals were finding their task more difficult—due to the conflict—than the less-than controls; whereas the greater-than experimentals were finding their job easier than the greater-than controls. Presumably the cooperation between the two dimensions made it easier for them to make up their mind, though more erroneously.

This discrimination hypothesis clearly implies that both sensory and social variables obey the same law. There is no reason to expect any difference between the two except that a sensory variable possesses an immediacy that is lacking in something like value, which is more abstract and conceptual; thus while we expect the form of the law to be the same it seems likely that sensory variables will be more effective.

I know of little evidence supporting this last prediction except for an inadequate earlier study of my own. In this earlier study, the subjects were to judge the number of dots on a card; and the contaminating variable was the size of the card upon which the dots appeared. Due to various difficulties only 35 subjects were tested, and the whole treated as a pilot experiment. Nevertheless it turns out that the equation

$$A'_j = 1.740 A_j + .051 Q_j - .687 \quad [7]$$

gives a rather good fit to these data. The material does not warrant a significance test, but it is interesting that the d here is about 10 times as large as our previous d . Incidentally, the large k here is a result of the fact that the square which was regarded as least numerous by the experimental group was a different card from the one regarded as least numerous by the controls.

F. SUMMARY

It is argued that the experimental paradigm for perceptual-accentuation studies should be one of psychophysical scaling in which the objects involved

are scaled on a variable, scaled again on the same dimension when the judgments were made in the presence of a disturber, and the disturber variable itself scaled. The disturber variable is ordinarily value, but need not be. The object is to discover the function relating the three sets of scale values. In the present study the scale values gotten in the presence of the disturber variable were found to be a linear transformation of the scale values gotten in the absence of the disturber variable plus an increment which is a function of the extremeness of the object in terms of the scale values on the disturber variable. A hypothesis explaining these results in terms of a discrimination process is offered.

REFERENCES

1. BIRKHOFF, G. D. *Aesthetic Measure*. Cambridge: Harvard Univ. Press, 1933.
2. BRUNER, J. S., & GOODMAN, C. C. Value and need as organizing factors in perception. *J. Abn. & Soc. Psychol.*, 1947, **42**, 33-44.
3. GUILFORD, J. P. *Psychometric Methods* (2nd ed.) New York: McGraw-Hill, 1954.
4. ROSS, R. T. Optimal orders in the method of paired comparisons. *J. Exper. Psychol.*, 1939, **25**, 414-424.

Department of Psychology
University of Miami
Coral Gables 46, Florida

ROLE-DEVIANT RESPONDENT SETS AND RESULTING BIAS, THEIR DETECTION AND CONTROL IN THE SURVEY INTERVIEW*¹

Family Service of the Cincinnati Area and University of Cincinnati

MERTON S. KRAUSE

A. INTRODUCTION

Research interviews are conducted for the collection of data relevant to certain variables of interest to the investigator. The data are coded, translated, or interpreted into estimated values for these variables and so mediate between the phenomenal referents of the variables and their value estimates. Whatever else, other than the relevant phenomena, influences the data and the estimates derived from the data constitutes a possible source of bias in the estimates. Since the critical manifestation of interview-data validity is in the value estimates for the variables of interest, data bias is of ultimate practical significance in the effect it has on the final statistics [i.e., as "net error" (32, pp. 228-229)].

Research-interview data proper are, of course, collected in the form of answers to questions. *Whatever tends to divert a respondent's answers from being maximally responsive, recordable, honest, understandable, considered, and complete, is a possible source of bias.*² We shall refer to this class of bias sources as respondent-bias sources even if induced by the interviewer or other "causes." There are also question-bias sources, involving the systematic elicitation of irrelevant information in the answers (45, 54) and interviewer-bias sources involving surmise, respondent selection, inadequate probing, and recording errors (32). We shall concern ourselves in this paper solely with the analysis of the detection and control of respondent bias sources.

* Received in the Editorial Office, Provincetown, Massachusetts, on April 15, 1964. Copyright, 1965, by The Journal Press.

¹ This analysis was supported under Public Health Service Research Grant MH 05031 from the National Institute of Mental Health. Bibliographic assistance was rendered by John Fine and constructive criticism by Mrs. Carroll Krause, Robert Ryder, and Mrs. Carolyn Winget.

² An analogous problem arises in experiments with human subjects which is discussed by M. T. Orne (53). Our numerous citations of the "objective" psychological testing literature testify to the existence of the same problem in that situation, and Lindzey (44, pp. 160-170) reviews it in the projective-testing area. We shall, however, restrict our discussion to the voluntary and putatively direct type of interview or test situation (8); though our citations range more widely.

B. RESPONDENT SET

Between the statement of a question by the interviewer and the statement of an answer by the respondent, a great deal of "mental activity" may take place in the respondent. He must grasp the meaning of the question (and we shall assume he does), select and assemble the relevant knowledge available through perception or memory, formulate a message to achieve whatever effects he intends, and verbalize this message as his answer. The limits of his perception, memory, and experience—as well as of his conceptual and verbal skills—limit the validity of his answer (e.g., 59, 60). Barring the influence of interviewer-bias or respondent-bias sources, these limits should manifest themselves in easily detectable ways: i.e., as "Don't Know," "Uncertain," unobtainable, incomplete, insufficient, unintelligible, unrecordable, or uncodable answers. With the exception of scoring procedures that make use of such special answers, these limits do not contribute to bias directly but rather only indirectly by attenuating in unknown ways the sample of respondents upon which the estimates are based (just as do refusals to be interviewed and "not-at-home" 's). This attenuation or sampling-bias problem, however, is beyond the scope of our analysis.

It is in the respondent's formulation of his answer, constrained by certain implicit goals—which the answer is (partially) fashioned to accomplish—that the respondent-bias sources operate. *When these bias sources are "ulterior-motivational" determinants* (i.e., not directed *only* toward responsive, recordable, complete, thoughtful, understandable and honest answering), *we shall refer to them as role-deviant "respondent sets."* They are not merely limits, for they may lead respondents willfully to give dishonest, misrepresentative, or irrelevant answers. These sets may produce answers that are more or less than mere responses to the specific explicit questions asked in that they use the answer form as a vehicle for their own ends.³ In a "psychology of the other one" (i.e., behavioral psychology) sets, as we shall use the term, are dispositional constructs akin to motives (39) but generally restricted in reference to behavior in the interview or test situation.

What sets are operative in an interview depends upon the nature of the respondent's goals for the attainment of which he perceives the interview

³ This general line of analysis was initiated, we believe, by J. W. Getzels (26); but also see Vidich and Bensman (62, pp. 21-24). The assignment of set to hypothetical-construct status is consonant with Cronbach's usage (12), but the "functional" or motivational slant seems to differ from his emphasis on phenotypic-response consistencies: i.e. its habit character. It seems wholly consonant with Meehl and Hathaway's discussion (47). We have presented elsewhere a discussion of role deviance and its sources in another context (38).

situation as being instrumental. His view of the situation activates or makes salient certain trans-situational motives (1, pp. 598-602; 46) that manifest themselves as answer-constraining sets, as well as in other ways. But the nature of the possible respondent bias depends more immediately on the specific sets operant during the interview than on their, possibly multiple, motivational sources. Therefore, bias control requires prior set detection for which respondent-personality assessment cannot be substituted even if such prior assessment were feasible (e.g., 24).

C. SET-DETECTION METHODS

There are at least six distinct methods of set detection. We shall start with the simplest and work our way through the more subtle and complex.

The respondent's goals in the interview situation may be, at least in part, quite overt. The more these goals manifest themselves outside of merely constraining the content of responsive answers (i.e., in byplay), the easier their detection. When the respondent is either willing to admit explicitly his ulterior goals or unable to control (or recognize) his byplay that expresses them, then this byplay readily lends itself to set detection. Thus, self-righteous uncooperativeness, fatigue, or deference to the interviewer's opinions might well be manifest in byplay (including questioning the interviewer) because they are not likely to be frustrated by admission, and the interview situation readily permits their expression in ways independent of the content of the answers. Its ready expressibility in byplay favors such manifestation of a set like wanting to obtain the interviewer's approval of oneself though it may be frustrated by admission. Neither expressibility nor admissibility favors the spontaneous expression in byplay of a set based, for example, on a fear of disclosing the fact of a criminal conviction (but the respondent's responsive byplay to a relevant question could serve to betray the set's existence). In sum, *continual byplay expressive of sets concerning gross and constant aspects of the situation*, rather than merely specific questions, should generally provide behavioral grounds for inferring respondent set.

A respondent's answers may be constrained by sets with regard to their understandability, recordability, responsiveness, completeness, thoughtfulness, or honesty. The first four of these answer variables, when deficient and not corrected by probing, contribute only to indirect (i.e., sampling) bias. Possibly less obvious and therefore less amenable to correction, thoughtless or unconsidered answers, which can occur in forms such as "I don't know" or "I'm not sure" and may be merely incomplete or undetailed answers formally satisfying the question, can introduce direct respondent bias. Where-

as unresponsive or formally incomplete answers are tantamount to admissions of a set in contravention to the proper respondent role, many thoughtless and dishonest answers do not betray themselves or the set that they manifest. Mere thoughtlessness may be disclosed by probing, but dishonesty often is hidden by a respondent for, when it is admitted, it is no more effective than bald unresponsiveness in achieving the respondent's goals and may have unfavorable side effects as well (e.g., loss of the interviewer's trust and approval). Therefore, a respondent's dishonesty, when intentional, probably is the most directly biasing, least discernible, and least remediable expression of any set.

The content of a respondent's answers can be the basis of a number of methods for the detection of sets concerning gross and constant aspects of the interview situation. All these are based upon the detection of dishonest answers and, just as in the case of byplay, dishonesty is contingent upon the perceived or apparent opportunity to advance the ends of its determining set. Thus, in order to seek the approval of his interviewer through deference to the latter's opinions, a respondent must believe he knows what these opinions are relative to the questions asked of him (42). In order to lend himself status in his interviewer's eyes, a respondent must believe he knows how to answer the interview questions to this end (65). The *consistency of his answers as efforts toward attaining a constant goal* suggest that a respondent has a set that determines this consistency [see (19, pp. 80-83)]. There are at least two empirical ways to determine what pattern of answers is to be associated with a given set. First, have an independent or a matched sample from the respondent population rate the various possible answers to each question with regard to their probable effectiveness as a means to the satisfaction of a given set under the projected survey conditions; e.g., rate the social desirability of the various answers (19). Second, have the same kind of sample attempt to answer the questions in such a way as most likely to attain satisfaction of the given set (and, therefore, not give away their intentions when this would jeopardize set satisfaction). This could be accomplished in a role-playing (faking) or real-life situation (16, 58). However, in subsequently applying such a group of studied questions, it may have to be supplemented by other questions beyond those essential to the primary research requirements in order that set-indicative answer patterns can clearly manifest themselves unconfounded with the substantive interview variables [see (11, pp. 152-153), 13, 30, 68, p. 233)]. If these answer patterns are both discernible and interpretable by a respondent, they will be inappropriate for the detection of his inadmissible sets.

Because it is possible that a series of honest, set-independent answers can also conform to the answer pattern of some set pattern (*cf.* 50) or that this latter has been imperfectly specified [discussed later under Bias Control through "answer balancing"], we must emphasize that answer patterns are not conclusive evidence in any individual respondent's case. Confirming by-play, of course, contributes to the evidence for the existence of a set and so do certain detection devices (in the interview) that reveal as such the respondent's efforts toward set enactment.

These devices are of at least three general types. The first type involves questions the possible answers to which have a considerable range of probable effectiveness and clear identifiability as efforts toward the satisfaction of some given set. The identifiability derives from *the unlikelihood of certain indicator answers being given, except for that set*. The independently determined falsity of the indicator answer or score for this respondent [e.g., a "sleeper question" (32, p. 227)], its rarity in the population (25, p. 162), or its implausibility [e.g., "I never worry" (47, pp. 537-538)] may be used as evidence of this unlikelihood. If memory errors, sampling error, and errors of interpretation are discounted (and if alternative sets can be ignored), this first type of detection device could provide sufficient though not necessary evidence for set detection. Since respondents may have idiosyncratic conceptions of effective set enactment or hold views of answer rarity and plausibility comparable to those of the interview coder (27, pp. 17-18), the absence of indicator answers from this device cannot be held to be conclusive evidence for the absence of a set.

The second type of detection device involves combinations of *questions that can yield a certain internally contradictory pattern of answers only when a given set is active* (42). Thus essentially the same question could be asked worded positively (i.e., "You have . . . ?"); later, negatively (i.e., "You haven't . . . ?") to detect an acquiescence or deference set (28; 9, p. 328). Of course, a respondent's memory of what he said before and his skill at qualifying answers can help him conceal his set; so the absence of an indicator pattern is, again, inconclusive. This type of device is prone to alert the more intelligent respondents to our detection efforts and so itself elicit sets. Although subtlety of question wording can disguise the device, it also tends to blur the possible internal contradictions of the respondent's answers.

The third type of detection device involves *questions relevant to two or more sets that conflict for each possible unqualified answer*. Thus, "Like many people, you've sometime thought about seeing a psychiatrist haven't you?" could pit an acquiescence set against a social-desirability set for many

people. The promise in this device lies in the respondent's self-revelatory qualifications in his extended answer which may disclose but need not disclose his set [for some indicator-behavior description *see* (35, pp. 47-49)].

The occurrence of spontaneous set-disclosing byplay, the matching of overall answer patterns to those produced as manifestations of known sets, and the appearance of indicator answers to built-in detector questions are means of set detection that at most require sharp observation and reliable recognition by the interviewers. There is another means that requires considerably more: namely, controlled variations in reaction to the respondent's answers and byplay. This is not a departure from the model interviewer's invariably patient, accepting, moderately friendly, always interested, and neither judgmentally reactive nor content-directive manner. The object of this reactive method is to give the respondent a *response to his (suspected) set enactment that is surprisingly incongruent with his expectation of its results*, but that does not impugn the interviewer's appearance of sincerity nor the rest of his role characteristics. An "unwittingly" but plausibly erroneous summarization of an answer contradictory to the client's apparent set (e.g., ". . . and you make almost \$10,000 a year before taxes and . . ." to a respondent suspected of a social-desirability set who reported ". . . over 10 . . .") is one tactic (17, pp. 480-481). If the respondent displays emotional involvement or an excessively effortful correction of the interviewer's "error," these may be taken as evidence of a set, though not of bias itself; while the absence of such effects is wholly inconclusive. (The probing of possibly set-biased answers may also evoke this sort of evidence.) However, the possible side effects of such emotional arousal make the use of this method advisable only infrequently or at the end of interviews.

From the preceding discussion, the crucial role of hypotheses about the nature of the sets to be encountered should be evident. The identification of set-pure answer patterns, detector-question design, and interviewer training in reaction tactics clearly presuppose such hypotheses; and alertness to byplay patterns and to reactions to probing would benefit from having ideas of what to look for.

D. RESPONDENT SETS AND THEIR CONTROL

Although persuasive empirical evidence of what motives manifest themselves in a given respondent's interview behavior might be impracticably difficult to assemble in a purely inductive manner (*see* 39), an a priori list of probably influential respondent goals can be assembled.⁴ Such a list, however

⁴ They resemble some of the factors extracted in recent studies of derived response-set scores (7), social-desirability ratings (48), and response-set scales (3).

generated, would at least enable us to attempt some rational set detection and would be amenable to empirical testing and amendment (41, p. 194).

We believe that most American respondents have, though to different degrees, at least the following goals in a research-interview situation: (*a*) to appear polite, (*b*) to appear honest, (*c*) to express one's side, opinions (or even mood) and to be understood, (*d*) to avoid embarrassment or loss of face, (*e*) to avoid getting into trouble through what one says, and (*f*) to avoid excessive effort or emotional upset. The first three of the foregoing are consonant with ideal respondent role enactment: responsive, understandable, recordable, formally complete, considered, and honest answers to all questions. However, they may also manifest themselves in (*a*) apparently socially desirable and acquiescent answers, (*b*) apparently socially desirable and lie-concealing consistent answers, or (*c*) nonresponsively self-assertive answers respectively. The last three goals—avoidance goals—are probable sources of only social-desirability and personal-desirability (57) biased answers. They tend to influence answers with respect to the latter's (*d*) probable effects on the interviewer or others present, (*e*) salient others in general who may get access to the respondent's answers and use them against him, and (*f*) the respondent himself respectively. These dispositions to achieve goals which may conflict with ideal-respondent role enactment are the sets upon which we shall base our discussion of specific set detection and control.

1. *Politeness*

The disinclination of a respondent to discountenance or displease his interviewer will depend in the way it affects his behavior upon what actions he believes would cause this to happen [*see* (32, pp. 162-167)]. The more "flank exposed" by expressing or implying attitudes, even inadvertently (*see* 31), liable to contradiction by complete and honest answers, the more vulnerable the interviewer makes himself to bias from this quarter.⁵ Mere ambiguity, however, may induce the respondent to react solely on his own needs and suspicions (5). Therefore, the best initial tactic would seem to be a plausible definition of the interviewer's role and evidence of his durability and neutrality (15) as well as the avoidance of excessive: i.e., seductive, rapport [*see* (32, pp. 46-52)]. Prefaces to questions can be used to reinforce these points too (e.g., "some people I've talked to have said . . . while others tell me . . ."). When the interviewer suspects he has received a

⁵ Even neutrally stated leading questions may induce acquiescence [*see* (20)]. For a general review of studies relevant to influencing interviewees see Blake and Mouton (4).

politeness-biased answer, his probing should manifest these same points. Byplay and the effects of early detector questions may originally alert him to the danger.

2. *Honesty*

To appear honest a respondent must answer in a consistent manner, avoid or explain any evident reasons for being dishonest, and reject any implications of dishonesty. This means that should the respondent suspect the interviewer of doubting his honesty, he may attempt to preserve appearances (and he will tend to avoid arousing such doubt). The bias problem here is mainly one of spurious, lie-concealing consistency in a respondent's answers. This consistency itself, especially if reinforced with byplay cues, may alert the interviewer; but he must not give any indication of being suspicious or of having caught inconsistencies during his initial presentation of the interview material (66, p. 10 ff.). Confrontations are less disruptive of the interview as afterthoughts while "looking over" the completed forms, whether made most directly or in the form of probes of indicator answers to detector questions. However, they may be prerequisite to rapport with respondents who are "conning" the interviewer (36, pp. 55-56).

3. *Self-Expression*

The respondent's set to have his say would not seem to be a source of direct bias but rather a source of possible sampling bias in that the interviewer fails to get a full complement of answers from such respondents. Detection of this set is straightforward, but its control may be a trying matter (e.g., 60). A clear introduction by the interviewer and an explicit agreement as to his role by the respondent is a useful precondition to using reminders of the interview's purpose as an initial control tactic. When reminders fail, the interviewer has little choice but to gratify the respondent, interject what questions he can in a responsive manner, and patiently "wait out" the respondent in order to complete the interview.

When the self-expression is that of mood, it should be evident in the respondent's byplay; but it also may affect content—e.g., it may make him overly critical or pessimistic on attitude questions [see (35, pp. 89-99)]. The best control is probably to arrange a reinterview.

4. *Saving Face*

A respondent's efforts to achieve or maintain a sense of respect and acceptance from the interviewer may lend themselves to bias (32, pp. 167-169, 201-

203, and 311-312). They presuppose, however, some belief (whether explicit or vague and inchoate) about what the interviewer respects or rejects. Since this belief may be false or unknown to the interviewer (especially when he differs in class or culture from the respondent), the respondent's pattern of answers and byplay may be unrevealing, and reactive methods of detection are contingent upon clear suspicions. The respondent's possible concerns with social status (as reflected in the interviewer's clothing, manner and accent, language, or auspices), morality, political identity, mental health, patriotism, professional or local loyalties can be prepared for and can reasonably be expected. The general methods of detection are those we have discussed, but there are also certain general methods of control.

We should, first, attempt to render as many as possible of the substantive grounds for face saving, which seem likely sources of bias, irrelevant to the interview. Auspices unsuspected of having any strong answer preferences—e.g., not a church group for a survey on religious practices [*see* (32, pp. 185-190)], introductions and question sequences unevocative of ideas of interested reference groups for the respondent (52), interviewer's apparent class and sub-cultural "identity" neutral relative to the questions—these are optimal; but they presuppose considerable knowledge about the respondent population and that it is homogeneous.

Second, and equally difficult, an effectively neutral and convincing sponsor, purpose, and interviewer identity must be created to prevent biasing assumptions or projections by the respondents. Observed trial interviews are necessary to detect and eliminate obvious set-evoking cues from the interviewers.

Third, the equal acceptability to the interviewer of all answers ought to be demonstrated to the respondent. Precoded answers may be presented in the question itself as a set of equally reasonable alternatives. When a free answer is sought, alternatives can be demonstrated with probes. Introducing these with, e.g., "Several people have told me . . . have you run into this sort of thing?" or using the respondent's language well (15) can serve to display the interviewer's neutrality.

Finally, the interviewer must help the respondent establish an identity that he will not need dishonesty or evasion to maintain; thus it may be useful to appear comfortable in his home during the interview, if economic matters are to be discussed. Again, an opening series of relevant biographical questions may be helpful in preventing the respondent from attempting to create a false impression about his educational experience (say, by out-flanking him before he starts). It can also give the interviewer an opportunity to display his acceptance of and respect for the respondent as he really is.

5. *Foresight*

A prudent respondent will avoid harming himself by what he says (e.g., 32, pp. 146-149, 159-162 and 169-170). He must assess the probability that the wrong persons will be informed of his remarks and the degree of probable harm he would then suffer. This assessment presumably guides his disclosures. [Hyman's discussion of respondent anonymity is relevant here (32, pp. 182-185; also *see* 18).] An interviewer usually can hope only to influence the respondent's expectations about what persons will have access to the data or research reports. If the auspices, introduction, and interviewer's identity and manner do not seem sufficient to assuage a respondent's fears, there are at least two supplementary tactics available. The first and better tactic is to encourage respondents to reject questions they would prefer not to answer. This encouragement should come after some rapport has been established and some satisfactory respondent role enactment has occurred (and before any "dangerous" questions) in order to maximize *its* likelihood of being believed. Refusals to answer are obvious set indicators, do not yield direct bias, allow the interviewer to try alternate questions, may be discussed to explore the respondent's set or influence it, and (if the answer is not obvious from the mere fact of refusal, which would tend to vitiate this first tactic) reduce the danger of bias emanating from efforts to answer consistently. The second tactic calls for standard set detection, especially answer patterns (when the interviewer suspects who it is the respondent fears) or detector questions.

6. *Motivation To Enact the Respondent Role*

If the respondent openly displays or admits his unwillingness to be interviewed, only sampling bias—due to his incomplete or refused answers—is likely to result. If his resistance is covert, however, direct bias from dishonestly casual or mischievous or from unconsidered answers may result. In the former case, the respondent's overt disinterest, preoccupation, intellectual objections, or fatigue should probably be explicitly recognized (in as face-saving a manner as possible) and openly accepted as reasonable by the interviewer in a calm, polite, understanding, and frank manner. The subsequent discussion could then serve to motivate the respondent to enact his proper role, arrange another time, or terminate the interview.

The latter case, that of the ambivalent respondent, is quite another matter. He wishes to avoid or escape the interview, but he also wishes to go on with it or to leave the impression of wanting to go on with it. Thus, he either cannot or will not easily be induced to admit any desire to stop the inter-

viewing. He gives signs of strain, which may be exacerbated by any attempts to focus attention on the strain or to discuss it. These signs will appear as signs of tension, anxiety, distractibility, anger, or mood changes, in the onset of problems in hearing, comprehension, or expression. When the respondent's ambivalence concerns gross and constant aspects of the situation, such as the interviewer, the fact of being interviewed, or the general topic it will be difficult to distinguish it from possible alternative sources of strain, such as chronic states of "intrapsychic stress" (neuroses, for example) or current acute life problems. If this background level of strain does not diminish, regardless of the situation's increased familiarity with time, the interviewer's accepting and understanding manner, and the respondent's interest in the questions, there is little the interviewer can do. The need for ego-supportive measures beyond those described should be taken as a signal to extricate oneself from the role of research interviewer gracefully and without offending or further straining the respondent. The application of such measures should also be noted as a potential elicitor of politeness-set effects or, if too obvious, face-saving set effects.

When the respondent's signs of strain are transitory or variable, it suggests that his ambivalence concerns specific questions rather than gross and constant aspects of the situation, assuming that all set-evocative byplay on the part of the interviewer is absent. In this case the interviewer must exercise considerable skill relative to these questions in avoiding probing too deeply, varying the question order, tolerating diversions and free answers even when irrelevant or redundant, reducing the degree of directness or structure in the questions asked, avoiding inferential or interpretive restatements, and still working toward the goal of the interview as far as seem advisable. Such liberties on the interviewer's part may increase the danger of question and interviewer-bias sources. Also they presuppose exceedingly alert and sensitive interviewers, who have been trained to watch the effects of their questions as well as listen to the answers they elicit.

E. BIAS DETECTION

There is only one means of detecting bias in any given data or final statistics, and that is to compare them with less-biased data or statistics. When data on the same variables have been independently collected in two or more ways from the same respondents, one of which is believed to yield less-biased results than the others, bias detection is possible (32, pp. 226-227; 66, p. 48-49); otherwise it is not possible. Criterion or control data are selected as such because in all regards that are believed to be susceptible to bias they

either seem inherently more valid on the basis of rational argument or they are believed to resemble more closely another class of data that provides a more valid criterion were the other data available. Thus, bias detection requires that we have available comparable data or statistics of greater validity than the data or statistics at issue. When one questions respondents about matters of recorded fact, such as age, car ownership, or public assistance, it is probably easily demonstrable by rational argument that the official records are more valid than the respondent's answers are likely to be [*see* (65, p. 381)]. However, the subjects of many surveys are not matters of recorded fact and so are not amenable to such straightforward bias-detection comparisons.

In the area of attitude, opinion, or knowledge surveys, the criterion data are those which have been collected under different respondent-set conditions than the data at issue. (We are, of course, still slighting both interviewer-bias and question-bias sources.) If one is persuaded of a lower probability or degree of bias under the criterion conditions, he can prepare a correction factor to be applied to statistics derived from data gathered under the experimental (usually the more convenient or common) conditions. If one is not so persuaded, he is bound at least to include warnings in his reports when the conditions compared yield significantly different statistical values. His persuasion must derive from arguments concerning, *inter alia*, matters of set detection and control. The detection of probably biasing sets in the "experimental" conditions not detected in the criterion conditions, the absence of any other biasing sets in the criterion conditions, and a significant difference in final statistics between the two conditions (which can be traced to differences in answers that should reasonably have been affected by the set-control techniques used in the collection of the criterion data)—this is the evidence upon which arguments of comparative validity must rest.

F. BIAS CONTROL

The biasing effects of respondent sets can be controlled in several ways. Perhaps the most obvious way is to attempt to increase the respondent's motivation to answer the interviewer's questions responsively, recordably, understandably, completely, thoughtfully, and honestly, so that competing sets are frustrated. To this end we strive to find some way to elicit motivation in the respondent which we expect would induce him to enact his proper role: adopt auspices with which he identifies, project uses for *valid* data which would benefit or please him, bring appropriate membership of reference groups to his mind, reinforce his role enactment (22), inspire empathy with the interviewer's task of getting unbiased answers, etc. [*see* (34, pp. 65-

91)]. This means is the foundation of bias control. It is direct and preventive, necessary though not always sufficient. There are, however, at least six other, supplementary means of bias control possible (*cf.* 32, p. 276).

First, *set control*, which has been discussed, is intended to decrease the degree to which role-deviant sets constrain the respondent's answers and thereby prevent the occurrence of biased answers (32, pp. 308-312). It works by inhibiting, harmlessly satisfying, or avoiding eliciting motives in the respondent competitive with his proper role enactment (34, pp. 193-201).

Second, *set counterbalancing*, which acts to cancel the expected biasing effects of one set and the means of satisfying it by eliciting another set or means of set satisfaction of equal and opposite expected bias and so *prevent* bias [*see* (63) for a special case]. Thus, an apparent set to be polite, which may be manifest by exclusively favorable answers regarding the interviewer's auspices, might be balanced by a slightly critical comment from the interviewer about his sponsoring organization. Such a remark could serve to check the tendency toward positive bias in answers concerning this organization. Again, a respondent's tendency to restrain his criticism might be countered by probes challenging his perspicacity: "Several people had some strong things to say about the wait for an appointment, what was your impression of it?" If these are made in a calm, unprovocative manner but with the evident expectation that the respondent ought to have some impressions, they may elicit a politeness or face-saving set contrary in effect to his prudent caution. Counterbalancing clearly requires that the set ostensibly biasing the respondent's answers is known because the design of appropriate questions, probes, remarks, and introductions is contingent upon such knowledge. Pitting two oppositely biasing sets against each other is just what we proposed for the third type of set-detector question, but now we are interested in the respondent's final answer rather than in his process of answering. Naturally, the likelihood of perfect counter balancing is rather remote; so the best result we should strive for is to have the respondent compromise on a less-biased answer than if his natural set were unopposed. Care must be taken not to overcorrect his probable bias. No matter how weak is the conflict we induce in the respondent, however, it may act (*a*) to lengthen the interview through increasing decision time (e.g., 2); (*b*) to increase the respondent's strain (43, pp. 260-269), (*c*) to decrease the test-retest reliability of the answer (35, pp. 15-67), and (*d*) to cause fluctuation in direction of imbalance from question to question.

Third, *answer balancing* aims to equate the degree of set satisfaction for a given set in the several possible alternative answers to a question (19, pp.

59-71) or in the answers to several questions directed at the same information (e.g., 23). If successful this would render that set incapable of biasing respondent's answers to the question and so prevent bias rather than merely correct it. To achieve this, however, the alternative answers must be completely indistinguishable as to their degree of the given set's satisfaction. Therefore, some means of selecting a group of balanced alternatives must be used. Edwards (19, pp. 1-31) describes one means of approximating such a result for class-determined or subculture-determined sets by using an independent random sample of the respondent population to scale the alternatives with regard to their degree of satisfaction of probably relevant, common, and influential sets.⁶ If this scaling can be done and is relatively free of bias, if the population of interest is homogeneous in natural set, if the reliability over subjects of the scale-value assignments is high,⁷ if the bias contributions of the various relevant sets are additive or a full factorial design was used in the scaling study [see (33, pp. 247-249)], and if the contributions of other sources of respondent bias than the specific sets covered in the scaling study are negligible, then and only then could a reasonably near answer balance be assured. However, anything short of perfect balance would probably continue to yield considerable bias (e.g., 10; 21) and only result in making the set-dominated respondent's forced choice more difficult (19, pp. 59-60) which may tend to induce conflict in these respondents (19, p. 60; 40).

The degree of balance achieved in a group of alternative answers with regard to their satisfaction of class-determined or subculturally determined sets can be tested for certain respondent populations by comparing the answers of a random sample of respondents motivated to enact a given set (i.e., to "fake") with those of an independent random sample *not* so motivated (19, pp. 55-58; 58). If the popularity of the various alternative answers to a question differs significantly between the two samples, these answers are not balanced relative to the given set for this respondent population (e.g., 6). Of course, population heterogeneity with regard to natural set, high item variances from other sources, or considerable popularity of the indicator answers under the control condition (68, pp. 230-232) would necessitate large subject samples to minimize the likelihood of erroneously finding a condition of balance to exist.

A logical extreme of answer balancing is the use of nonloaded or evenly

⁶ The social desirability variable studied by Edwards may actually consist of several variables (e.g., 6, 19, 48, 57) and be overshadowed by idiosyncratic personal desirability (e.g., 29).

⁷ This may well be lowered by variance due to differences in interview settings [see (37) and (55)].

loaded questions (54, pp. 177-202): i.e., of set-irrelevant, subtle, or neutral questions (67). These are tantamount to answer balancing at relatively low set-satisfaction (or dissatisfaction) levels, though their development primarily involves problems of question rather than answer selection. Empirical procedures like those described are properly included in the selection or testing of such questions.

Fourth, *intrainterview bias-elicitor balancing* involves using two or more questions in different forms or positions, with different introductions or asked in different ways which are designed to elicit the same information in a manner such that some combination of the answers is relatively unbiased (14). It is appropriate to the correction of bias in single interviews. There are two distinct models of bias-elicitor balancing: the probabilistic and the determinate. The former assumes, for example, that the biasing effects of the universe of sets are symmetrically distributed about zero, so that a random sample (with respect to set elicitation) of questions-in-context about the same information would tend to have a mean bias of zero (but for sampling error). The major obstacle in principle to realizing this model, if it is valid, is the drawing of such a sample, since the necessary universe of questions-in-context is merely a theoretical construct. However, the practical effects of asking several such questions in an interview (i.e., as itself a source of set elicitation) may be as serious an obstacle.

The determinate model, though also subject to the latter practical problem, has a different theoretical problem to face. It assumes that any set against which it is used can be reversed in direction of bias, though unchanged in degree, by some alteration of question-in-context. The simplest instance of this model is probably that designed to control intrainterview bias due to an acquiescence or politeness set (or its converse, an oppositional set). Thus, one might ask a respondent to express his agreement or disagreement to a pair of statements (separated in the interview by considerable other "neutral" material), such as (a) "It's hardly fair to bring children into the world the way things look for the future," and (b) "Children born today have a wonderful future to look forward to" (from 42). There are four possible answer patterns: (a) agree-disagree, which is pessimistic, (b) disagree-agree, which is optimistic, (c) agree-agree, which is acquiescent, and (d) disagree-disagree, which is oppositional. If single optimistic answers receive a unit score and single pessimistic answers receive a zero score, the sum of the residual (unbiased) answers⁸ would be zero, two, one, and one respectively. Thus,

⁸ If we have n questions phrased as (a) and n phrased as (b) and if a respondent's total agreements with (a)'s are denoted by X_{aa} , total agreement with (b)'s by X_{ba} ,

for the possible direct bias of either question alone, we can through set-elicitor balancing substitute the indirect or sampling bias of an unusable total response (as a score of one may be interpreted as indicating here). This means of bias control is, of course, homologous to the second type of set-detection device. It must be noted that with more questions than two, considerably more refined and complex, though not obviously preferable, scoring procedures are possible [see (30)].⁹

Fifth, *interinterview bias-elicitor balancing* can serve to minimize net bias in answer distributions for individual questions or groups of questions. This is possible when some situational, procedural, or interviewer variables are reliably divisible into subgroups or subranges of which two are suspected of eliciting biasing sets opposite in direction of bias. When both supposedly biasing subgroups are used, they may be balanced against each other.

If we assume that a question of interest has only two possible answers, A_1 and A_2 , that the proportion of respondents giving A_1 in the absence of any relevant set elicitors is π and the proportion giving A_2 under these conditions is $(1 - \pi)$, that the relative effectiveness of the set elicitor inducing naturally A_2 respondents to give A_1 is α_1 while the relative effectiveness of its opposite number favoring A_2 is α_2 , and that the random sample of respondents upon which α_1 operates is $1 \geq p_1 \geq 0$ of the total sample of respondents while α_2 operates on $p_2 = (1 - p_1)$, then the net bias will be zero if and only if

$$\frac{\alpha_2 \pi}{\alpha_1 (1 - \pi)} = \frac{p_1}{1 - p_1}.^{10}$$

total disagreement with (a)'s by X_{ad} , and total disagreement with (b)'s by X_{bd} , then $[X_{ad} + X_{ba}/X_{aa} + X_{ad} + X_{ba} + X_{bd}]$ will yield this sum score. If this score is significantly different from one-half, we have convincing evidence that something besides chance or the set at issue is involved in the respondent's pattern of answers. Unless other biasing factors are suspected, we may conclude that this something is—in our example—his attitude toward the future. A score of one-half due to entries in only one column of the X matrix,

	Agree	Disagree
a phrasing	X_{aa}	X_{ad}
b phrasing	X_{ba}	X_{bd}

evidences pure set determination of the respondent's answers. Sample estimates of the relative contributions of set and honesty to the obtained results are also possible (9, 49).

⁹ Although equal numbers of questions phrased on each side of the balance is not a necessary condition of using these procedures, it does (a) safeguard equally against either direction of bias and (b) provide the most conservative null hypothesis because the population proportion of highest sampling error is one-half.

¹⁰ Since the respondent bias in the estimate of π from the sample is $b_1 = a_1 p (1 - \pi) - a_2 (1 - p) \pi$ and that in $(1 - \pi)$ is $b_2 = -b_1$; if $b_1 = 0$, the equation is readily derived.

Even rough estimates of α_2/α_1 and $\pi/(1-\pi)$ can serve to reduce net bias (through the adjustment of p_1 and p_2) if they are in the correct direction from unity. In the absence of all reasonable grounds for estimates, $p_1 = .5$ has been recommended as the best bet because the maximum *possible* net bias, given some particular numerical difference between an estimate of π based on the respondents subject to α_1 and an estimate based on those exposed to α_2 , is less for this value of p than for any other (32, pp. 315-316; 51, pp. 186-188).

This procedure would be appropriate, for example, where a relatively wide range of discriminable social-class members were being used to interview middle-middle-class respondents, and it was suspected that these respondents would be biased in reporting their expenditures due to face saving tendencies with upper-middle-class interviewers and due to politeness with lower-middle-class interviewers. If the average expenditures for the sample respondents is the most important statistic to be obtained, this is the procedure of choice. If, on the other hand, certain cross-classifications are of primary interest, it is gross bias regardless of direction rather than net bias that should be minimized (32, pp. 317-318): i.e., Procedures 1 through 4 are then preferable [but see (61)].

Sixth, *statistical adjustment* presupposes bias detection and measurement, as do Procedures 2 through 5; but uses this information to correct rather than prevent bias. Given an observed respondent answer or score or sample statistic, we must know how much to adjust it in what direction. This depends, for our purposes in this paper, upon the following several factors: (a) respondent-set elicitation or activation, (b) bias opportunity, (c) bias propensity [see (56)] given (a) and (b), and (d) actual bias (i.e., the signed difference between observed and unbiased answer, score, or statistic). The last is all we really must know to make statistical adjustments, but usually it will be more feasible in practice to use combined measurements or estimates of the other factors. If we may suppose that these factors each assume values of one or zero for any specific respondent and are statistically independent, a simple example can be given.

If (a) 40 per cent of a respondent sample were, say, induced to try saving face on questions concerning income, (b) all respondents were asked whether or not their income was over a certain amount and for 75 per cent it was not, and (c) half of those who were motivated to save face and could positively bias their answers [i.e., $.5 \times (.4 \times .75) = 15$ per cent of the sample] did so, then knowledge of (a), (b), and (c) for each respondent would permit perfect bias correction for this source, for it is tantamount to recognition of every answer biased from this source and of the unbiased

answer for each. Any lesser degree of information¹¹ would make the correction of gross bias imperfect, while only precise sample values of (*a*), (*b*), and (*c*) would suffice to wholly correct net bias.¹² Thus, even in the simplest case the information required for gross (respondent-by-respondent) statistical adjustment for bias is unlikely ever to be available, and even that required for net (sample-estimate) adjustments can only be roughly estimated: most roughly for (*c*); least for (*a*)—if set detection is vigorously pursued.

G. CONCLUSIONS

Bias control is our ultimate objective, but in order to know how closely we have achieved it (and, for Procedures 5 and 6, even to attempt it) bias detection is required. Insofar as role-deviant respondent sets are the source of bias, their control can reduce bias and, on an experimental basis, provide the conditions for bias-detection studies. Set control, however, can only be evaluated and, on the basis of prior studies of the same respondent population, efficiently designed to the extent that set detection is effectively utilized. The remaining respondent-bias control procedures, the second through the sixth, also depend upon prior set-detection studies and therefore on effectively utilized set detection. Thus, the validity of survey-interview data and statistical estimates can be insured against respondent bias only by prior studies for set detection and then for comparisons of various relevant set and other respondent-bias control procedures. In this way *only* can an investigator systematically minimize the possible respondent bias in his attitude-survey, opinion-survey, or knowledge-survey findings.¹³ Unfortunately, he cannot be certain of having effectively avoided all respondent bias because some unanticipated sets may be operative and because the set-detection methods are prone to false-negative results. This limitation is not grounds for ignoring the preparatory research on respondent bias suggested, for such research is a prerequisite of interview-situation design, questionnaire design, interviewer selection and training plans, the survey's "advertising campaign plans," and the construction of statistical adjustments.

¹¹ It is obvious that observed answers, scores, or statistics cannot be properly adjusted on the basis of (*a*) or even a_1 information alone [as Webster (64) has proposed; see (50)].

¹² For the sample $ac = a_1$ and $b = \pi$, as a_1 and π were used in discussing Procedure 5. Therefore, given a_1 and π and $p = 1$, $b_1 = a_1 (1 - \pi)$, and so the estimate π is corrected by subtracting b_1 from it. In our example above, $\pi = \hat{\pi} - .15 = 25$ per cent [see (32, pp. 316-317)].

¹³ Except where he is repeatedly predicting a criterion (e.g., an election) and can develop appropriate item weights by multiple-regression techniques.

To develop a maximally valid survey of opinions or attitudes and so minimize the possible errors attributable to poor sampling procedures, question ambiguity, or uncontrolled respondent bias, one requires a research project that may be more demanding than the final survey.

REFERENCES

1. ATKINSON, J. W. Thematic apperceptive measurement of motives within the context of a theory of motivation. In J. W. Atkinson (Ed.), *Motives in Fantasy, Action, and Society*. New York: Van Nostrand, 1958.
2. BARKER, R. G. An experimental study of the resolution of conflict by children. In Q. McNemar & M. A. Merrill (Eds.), *Studies in Personality*. New York: McGraw-Hill, 1942.
3. BENDIG, A. W. A factor analysis of "social desirability," "defensiveness," "lie," and "acquiescence" scales. *J. Gen. Psychol.*, 1962, **66**, 129-136.
4. BLAKE, R. R., & MOUTON, J. S. The experimental investigation of interpersonal influence. In A. D. Biderman & H. Zimmer (Eds.), *The Manipulation of Human Behavior*. New York: Wiley, 1961.
5. BORDIN, E. S. Ambiguity as a therapeutic variable. *J. Consult. Psychol.*, 1955, **19**, 9-15.
6. BORISLOW, B. The Edwards Personal Preference Schedule (EPPS) and fakability. *J. Appl. Psychol.*, 1958, **42**, 22-27.
7. BROEN, W. E., JR., & WIRT, R. D. Varieties of response sets. *J. Consult. Psychol.*, 1958, **22**, 237-240.
8. CAMPBELL, D. T. A typology of tests, projective or otherwise. *J. Consult. Psychol.*, 1957, **21**, 207-210.
9. CHAPMAN, L. J., & BOCK, R. D. Components of variance due to acquiescence and content in the F scale measure of authoritarianism. *Psychol. Bull.*, 1958, **55**, 328-333.
10. CORAH, N. L., FELDMAN, M. J., COHEN, P. S., GRUEN, W., MEADOW, A., & RINGWALL, E. A. Social desirability as a variable in the Edwards Personal Preference Schedule. *J. Consult. Psychol.*, 1958, **22**, 70-72.
11. COUCH, A., & KENISTON, K. Yeasayers and naysayers: Agreeing response set as a personality variable. *J. Abn. & Soc. Psychol.*, 1960, **60**, 151-174.
12. CRONBACH, L. J. Response sets and test validity. *Educ. & Psychol. Meas.*, 1946, **6**, 475-494.
13. CROWNE, D. P., & MARLOWE, D. A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 1960, **24**, 349-354.
14. CURTIS, R. F., & JACKSON, E. F. Multiple indicators in survey research. *Amer. J. Soc.*, 1962, **68**, 195-204.
15. DEXTER, L. A. Role relationships and conceptions of neutrality in interviewing. *Amer. J. Soc.*, 1956, **62**, 153-157.
16. DICKEN, C. F. Simulated answer patterns on the EPPS. *J. Appl. Psychol.*, 1959, **43**, 372-378.
17. DOHRENWEND, B. J., & RICHARDSON, S. A. Directiveness and nondirectiveness in research interviewing. *Psychol. Bull.*, 1963, **60**, 475-485.
18. DUNNETTE, M. D., & HENEMAN, H. G., JR. Influence of scale administrator on employee attitude scale. *J. Appl. Psychol.*, 1956, **40**, 73-77.
19. EDWARDS, A. L. *The Social Desirability Variable in Personality Assessment and Research*. New York: Dryden, 1957.

20. ELLIOT, L. L. Effects of item construction and respondent aptitude on acquiescence. *Educ. & Psychol. Meas.*, 1961, **21**, 405-415.
21. FELDMAN, M. J., & CORAH, N. L. Social desirability and the forced choice method. *J. Consult. Psychol.*, 1960, **24**, 480-482.
22. FIELD, J. B. The effects of praise in a public opinion poll. *Pub. Opin. Quart.*, 1955, **19**, 85-91.
23. FORDYCE, W. E., & LAMPHERE, A. V. The control of social desirability in a structured Q-sort on dependency-independency. *Educ. & Psychol. Meas.*, 1960, **20**, 103-110.
24. FOSTER, R. J. Acquiescent response set as a measure of acquiescence. *J. Abn. & Soc. Psychol.*, 1961, **63**, 155-160.
25. FRICKE, B. G. Response set as a suppressor variable in the OIAS and MMPI. *J. Consult. Psychol.*, 1956, **20**, 161-169.
26. GETZELS, J. W. The question-answer process: A conceptualization and some derived hypotheses for empirical examination. *Pub. Opin. Quart.*, 1954, **18**, 80-91.
27. HANLEY, C. Social desirability and response bias in the MMPI. *J. Consult. Psychol.*, 1961, **25**, 13-20.
28. HARE, A. P. Interview responses: Personality or conformity. *Pub. Opin. Quart.*, 1960, **24**, 679-685.
29. HEILBRUN, A. B., & GOODSTEIN, L. P. Relationship between personal and social desirability sets and performance on the EPPS. *J. Appl. Psychol.*, 1959, **43**, 302-305.
30. HELMSTADTER, G. C. Procedures for obtaining separate set and content components of a test score. *Psychometrika*, 1957, **22**, 381.
31. HILDUM, D. C., & BROWN, R. W. Verbal reinforcement and interviewer bias. *J. Abn. & Soc. Psychol.*, 1956, **53**, 108-111.
32. HYMAN, H. Interviewing in Social Research. Chicago: Univ. Chicago, 1954.
33. JACKSON, D. N., & MESSICK, S. Content and style in personality assessment. *Psychol. Bull.*, 1958, **55**, 243-252.
34. KAHN, R. L., & CANNELL, C. F. The Dynamics of Interviewing. New York: Wiley, 1957.
35. KENDALL, P. Conflict and Mood. Glencoe: Free Press, 1954.
36. KINSEY, A. C., POMEROY, W. V., & MARTIN, C. E. Sexual Behavior in the Human Male. Philadelphia: Saunders, 1948.
37. KLETT, C. J. The social desirability stereotype in a hospital population. *J. Consult. Psychol.*, 1957, **21**, 419-421.
38. KRAUSE, M. S. A role-theoretic analysis of resistive behavior. Unpublished manuscript, 1962.
39. ———. Twelve propositions in the semantics of motivation. *J. Gen. Psychol.*, 1964, **70**, 331-339.
40. KRUG, R. E., & NORTHRUP, D. Judgment time for forced-choice adjective pairs. *J. Appl. Psychol.*, 1959, **43**, 407-410.
41. LANSING, J. B., GINSBURG, G. P., & BRAATON, K. An Investigation of Response Error. Urbana: Univ. Illinois, 1961.
42. LENSKI, G. E., & LEGGETT, J. C. Caste, class and deference in the research interview. *Amer. J. Soc.*, 1960, **65**, 463-467.
43. LEWIN, K. Field Theory in Social Science. New York: Harper, 1951.
44. LINDZEY, G. Projective Techniques and Cross-Cultural Research. New York: Appleton-Century-Crofts, 1961.
45. LITWAK, E. A classification of biased questions. *Amer. J. Soc.*, 1956, **62**, 182-186.

46. MARLOWE, D., & CROWNE, D. P. Social desirability and response to perceived situational demands. *J. Consult. Psychol.*, 1961, **25**, 109-115.
47. MEEHL, P. E., & HATHAWAY, S. R. The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *J. Appl. Psychol.*, 1946, **30**, 525-564.
48. MESSICK, S. Dimensions of social desirability. *J. Consult. Psychol.*, 1960, **24**, 279-287.
49. ———. Separate set and content scores for personality and attitude scales. *Educ. & Psychol. Meas.*, 1961, **21**, 915-923.
50. ———. Response style and content measures from personality inventories. *Educ. & Psychol. Meas.*, 1962, **22**, 41-56.
51. MOSTELLER, F. Correcting for interviewer bias. Appendix II in H. Cantril (Ed.), *Gauging Public Opinion*. Princeton: Princeton Univ., 1944.
52. NAKAMURA, C. Y. Salience of norms and order of questionnaire items. *J. Abn. & Soc. Psychol.*, 1959, **59**, 139-142.
53. ORNE, M. On the social psychology of the psychological experiment. *Amer. Psychologist*, 1962, **17**, 776-783.
54. PAYNE, S. L. *The Art of Asking Questions*. Princeton: Princeton Univ., 1951.
55. PHARES, E. J., & ROTTER, J. B. An effect of the situation on psychological testing. *J. Consult. Psychol.*, 1956, **20**, 291-293.
56. RIMLAND, B. Personality test faking: Expressed willingness to fake as affected by anonymity and instructional set. *Educ. & Psychol. Meas.*, 1962, **22**, 747-751.
57. ROSEN, E. Self-appraisal, personal desirability, and perceived social desirability of personality traits. *J. Abn. & Soc. Psychol.*, 1956, **52**, 151-158.
58. RUCH, F. L. A technique for detecting attempts to fake performance in a self-inventory type of personality test. In Q. McNemar & M. A. Merrill (Eds.), *Studies in Personality*. New York: McGraw-Hill, 1942.
59. SCHATZMAN, L., & STRAUSS, A. Social class and modes of communication. *Amer. J. Soc.*, 1955, **60**, 329-338.
60. STRAUSS, H., & SCHATZMAN, L. Cross-class interviewing. *Hum. Org.*, 1955, **14**, 28-31.
61. SUCHMAN, E. A. An analysis of "bias" in survey research. *Pub. Opin. Quart.*, 1962, **26**, 102-111.
62. VIDICH, A., & BENSMAN, J. The validity of field data. *Hum. Org.*, 1954, **13**, 20-27.
63. VOAS, R. B. A procedure for reducing the effects of slanting questionnaire responses toward social acceptability. *Educ. & Psychol. Meas.*, 1958, **18**, 337-345.
64. WEBSTER, H. Correcting personality scales for response sets or suppression effects. *Psychol. Bull.*, 1958, **55**, 62-64.
65. WEISS, D. J., & DAWIS, R. V. An objective validation of factual interview data. *J. Appl. Psychol.*, 1960, **44**, 381-385.
66. WELLMAN, F. L. *The Art of Cross-Examination* (4th ed.) New York: Garden City, 1948.
67. WIENER, D. R. Subtle and obvious keys for the Minnesota Multiphasic Personality Inventory. *J. Consult. Psychol.*, 1948, **12**, 164-170.
68. WIGGINS, J. S. Strategic, method, and stylistic variance in the MMPI. *Psychol. Bull.*, 1962, **59**, 224-242.

Institute for Juvenile Research
907 S. Wolcott Ave.
Chicago, Illinois 60612

BOOKS

Now that there is a special *APA* journal completely devoted to the publication of book reviews, it is no longer necessary that other journals emphasize such publication. It has always been our conviction that book reviews are a secondary order of publication unless they carry information that is equally important as the book. However, the publication of book titles is a very important service, and we shall continue to render that service.

In any given issue of this journal, we may continue to publish one or more book reviews, but we do not consider such publication a major function of this journal. In line with this policy, we can no longer pay for such manuscripts.

BOOKS RECENTLY RECEIVED

- ANDERSON, D. *The Psychology of Crime*. New York: Wiley, 1964. Pp. 358.
- ANGYAL, A. *Neurosis and Treatment: A Holistic Theory*. New York: Wiley, 1965. Pp. 328.
- BAROJA, J. C. *The World of the Witches*. Chicago: Univ. Chicago Press, 1965. Pp. 313.
- BERLYNE, D. E. *Structure and Direction in Thinking*. New York: Wiley, 1965. Pp. 378.
- BOWMAN, H. A. *Marriage for Moderns*. New York: McGraw-Hill, 1965. Pp. 709.
- BREER, P. N., & LOCKE, E. A. *Task Experience as a Source of Attitudes*. Homewood, Ill.: Dorsey, 1964. Pp. 280.
- CAPPON, D. *Toward an Understanding of Homosexuality*. Englewood Cliffs, N. J.: Prentice-Hall, 1965. Pp. 302.
- CRONBACH, L. J., & GLESER, O. C. *Psychological Tests and Personnel Decisions* (2nd ed.) Urbana: Univ. Illinois Press, 1965. Pp. 347.
- DAVIS, M. *Hope for the Childless Couple*. New York: McGraw-Hill, 1965. Pp. 197.
- DUVAL, R. *Adolescence d'Aujourd'hui*. Quebec: Les Presses de l'Universite Laval, 1964. Pp. 191.
- FARBER, S. M., MUSTACCHI, P., & WILSON, B. N. L., *Eds. The Family's Search for Survival*. New York: McGraw-Hill, 1965. Pp. 210.
- FARNSWORTH, P. R., *Ed. Annual Review of Psychology*. Palo Alto, Calif.: Annual Reviews, 1965. Pp. 571.
- FEIFEL, H., *Ed. The Meaning of Death*. New York: McGraw-Hill, 1965. Pp. 351.
- FREEMAN, R. ST. J., & FREEMAN, H. A. *Counseling: A Bibliography*. New York: Scarecrow Press, 1964. Pp. 986.
- FULTON, R., *Ed. Death and Identity*. New York: Wiley, 1965. Pp. 415.
- GARRISON, K. C. *Psychology of Adolescence* (6th ed.) Englewood Cliffs, N. J.: Prentice-Hall, 1965. Pp. 487.
- GREENFIELD, N. S., & LEWIS, W. C., *Eds. Psychoanalysis and Current Biological Thought*. Madison: Univ. Wisconsin Press, 1965. Pp. 380.
- GROSSMAN, C., & GROSSMAN, S. *The Wild Analyst*. New York: Braziller, 1965. Pp. 222.
- HAAS, K. *Understanding Ourselves and Others*. Englewood Cliffs, N.J.: Prentice-Hall, 1965. Pp. 370.

- HILLMAN, J. *Suicide and the Soul*. New York: Harper & Row, 1965. Pp. 191.
- HAUCK, M., & STEINKAMP, S. *Survey Reliability and Interviewer Competence*. Urbana: Univ. Illinois, 1964. Pp. 112.
- HULL, C. L. *A Behavior System*. New York: Wiley, 1964. Pp. 372.
- ILG, F. L., & AMES, L. B. *School Readiness*. New York: Harper & Row, 1965. Pp. 396.
- ISAACS, S. R. *India's Ex-Untouchables*. New York: John Day, 1965. Pp. 188.
- JOHNSON, R. C., & MEDINNUS, C. R. *Child Psychology: Behavior and Development*. New York: Wiley, 1965. Pp. 556.
- KERNHAUSER, A. *Mental Health of the Industrial Worker*. New York: Wiley, 1965. Pp. 354.
- LANGHOFF, P., Ed. *Models, Measurement and Marketing*. Englewood Cliffs, N. J.: Prentice-Hall, 1965. Pp. 216.
- LOPEZ, F. S. *Sociologia de la Accion*. Madrid: Instituto Balmes de Sociologia, 1964. Pp. 194.
- LORAND, S., & SCHNEER, H. I., Eds. *Adolescents: Psychoanalytic Approach to Problems and Therapy*. New York: Dell, 1965. Pp. 378.
- MACCOLL, S. H. *Structure and Development of Phenomenal Reality*. New York: Exposition Press, 1965. Pp. 222.
- MAIER, N. R. F. *Psychology in Industry* (3rd ed.) Boston: Houghton Mifflin, 1965. Pp. 718.
- MARTIN, J. M., & FITZPATRICK, J. P. *Delinquent Behavior*. New York: Random House, 1964. Pp. 210.
- MOK, P. P. *Pushbutton Parents and the Schools*. New York: Dell, 1964. Pp. 208.
- MORSE, W. C., CUTLER, R. L., & FINK, A. H. *Public School Classes for the Emotionally Handicapped: A Research Analysis*. Washington, D. C.: Council for Exceptional Children, 1964. Pp. 142.
- MUENSTERBERGER, W., & AXELRAD, S., Eds. *The Psychoanalytic Study of Society*. New York: Internat. Univ. Press, 1965. Pp. 408.
- MUSSEN, F. H., CONGER, J. J., & KAGAN, J., Eds. *Readings in Child Development and Personality*. New York: Harper & Row, 1965. Pp. 480.
- NEWCOMB, T. M., TURNER, R. H., & CONVERSE, P. E. *Social Psychology: The Study of Human Interaction*. New York: Holt, Rinehart & Winston, 1965. Pp. 591.
- PECK, R. F., & HAVIGHURST, R. J. *The Psychology of Character Development*. New York: Wiley, 1964. Pp. 267.
- PENLINGTON, N. *Canada and Imperialism, 1896-1899*. Toronto: Univ. Toronto Press, 1965. Pp. 288.
- PERLS, F., HEFFERLINE, R. F., & GOODMAN, P. *Gestalt Therapy*. New York: Dell, 1965. Pp. 470.
- PITTMAN, J., & PITTMAN, M. *Peaceful Coexistence: Its Theory and Practice in the Soviet Union*. New York: International Publishers, 1964. Pp. 156.
- REED, C. F., ALEXANDER, I. E., & TOMKINS, S. S., Eds. *Psychopathology*. New York: Wiley, 1964. Pp. 803.
- ROYCE, J. H., Ed. *Psychology and the Symbol*. New York: Random House, 1965. Pp. 117.
- RUBIN, S. *Psychiatry and Criminal Law*. Dobbs Ferry, N. Y.: Oceana, 1965. Pp. 219.
- RUITENBEEK, H. M. *The Individual and the Crowd*. New York: New American Library, 1965. Pp. 125.
- SCHILDER, F. *The Image and Appearance of the Human Body*. New York: Wiley, 1964. Pp. 353.
- TEEVAN, R. C., & BIRNEY, R. C. *Readings for Introductory Psychology*. New York: Harcourt, Brace & World, 1965. Pp. 504.
- VAN DEN HAAG, E. *Passion and Social Constraint*. New York: Dell, 1965. Pp. 368.

Journal of Genetic Psychology

Founded by G. Stanley Hall in 1891; reorganized by Carl Murchison in 1925

Quarterly, published in March, June, September, and December. Devoted to child behavior, animal behavior, and the problems of aging. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1891 at \$15.00 per volume, plus transportation.

Journal of General Psychology

Founded in 1927 by E. B. Titchener and Carl Murchison

Quarterly, published in January, April, July, and October. Devoted primarily to experimental, theoretical, clinical, and historical psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1927 at \$15.00 per volume, plus transportation.

Journal of Social Psychology

Founded in 1929 by John Dewey and Carl Murchison

Bimonthly, beginning with February. National, group, cultural, racial, and differential psychology. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Bimonthly, three volumes per year; per annum, \$30.00; per volume, \$15.00; single numbers, \$7.50. First number is dated February, 1930. Complete sets from 1930 at \$15.00 per volume, plus transportation.

Each author will receive 100 unbound reprints gratis. Additional reprints may be purchased at a nominal cost when the author returns page-proof.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

Club Rate for Five International Psychological Journals

THE JOURNAL OF GENETIC PSYCHOLOGY	PER ANNUM
(Quarterly, two volumes per year)	\$20.00
GENETIC PSYCHOLOGY MONOGRAPHS	
(Quarterly, two volumes per year)	\$20.00
THE JOURNAL OF GENERAL PSYCHOLOGY	
(Quarterly, two volumes per year)	\$20.00
THE JOURNAL OF SOCIAL PSYCHOLOGY	
(Bimonthly, three volumes per year)	\$30.00
THE JOURNAL OF PSYCHOLOGY	
(Bimonthly, three volumes per year)	\$30.00
TOTAL	\$120.00
TEACHERS AND STUDENTS EDITION, <i>Restricted Sale</i>	\$35.00

The *Teachers and Students Edition* is open to any individual psychologist, teacher, student, or physician but requires an agreement that the edition will not be offered for sale till after five years from publication and an agreement that the edition will not be substituted for an institutional subscription.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

Genetic Psychology Monographs

Founded in 1925 by Carl Murchison

Quarterly: February, May, August, November. Two volumes per year. Child behavior, animal behavior, and the problems of aging. Each number one or more complete researches. Manuscripts may be sent to any member of the Editorial Board, or may be sent directly to The Journal Press. All subscriptions and business communications should be sent directly to The Journal Press. Two volumes annually. Per annum, \$20.00; per volume, \$15.00; single numbers, \$7.50. Complete sets from 1926 at \$15.00 per volume, plus transportation.

NOTICE. Effective immediately we offer prior publication for certain selected monographs. Charges will be based on the announced rates for *The Journal of Psychology*. Our policy remains unchanged for regular monographs: Authors will be charged for cuts and expensive tables, but there will be no page charge for the first 50 pages. For all pages above 50 there will be a charge of \$10.00 per page. Two hundred reprints of each monograph will be supplied free of charge, but there will be nominal charges for such things as vitas, special covers, or special title pages.

The listing of monographs for 1926 through 1963 will be printed in the first half of each volume of all five journals. Monographs published after 1963 will be listed in the second half of each volume.

Genetic Psychology Monographs (continued)

VOLUME 69—January-June, 1964

1. Occupational and industrial therapy in the treatment of psychiatric patients: A controlled study of efficacy in an intensive-treatment institution—M. HUNDZIAK AND B. PASAMANICK
An analysis of Carl R. Rogers' theory of personality—M. S. KRAUSE
Personality aspects of teaching: A predictive study—H. ZIMILES, B. BIBER, W. RABINOWITZ, AND L. HAY
Situational and personal antecedents of incipient alienation: An exploratory study—W. H. JARRETT AND A. O. HALLER
2. Age changes in children's Mosaic responses from five to ten years—L. B. AMES AND F. L. ILG
A study of the interview process: The caseworker-client relationship—H. J. MEYER, E. F. BORGATTA, AND D. FANSHIEL
A study of introversion-extroversion as a personality dimension—R. M. ALLEN, H. M. RICHER, AND R. J. PLOTNIK
Psychological testing with Northwest Coast Alaskan Eskimos—C. E. PRESTON
A study in family pathology—M. B. HANS

VOLUME 70—July-December, 1964

1. Structure and function—H. GUTMAN
The Lowenfeld Mosaic Test: Norms for five-to-ten-year-old American public-school children and comparative study of three groups—L. B. AMES, F. L. ILG, AND J. AUGUST
Social and personality correlates of children's estimates of height—J. P. SHAFFER
The development of ethnic attitudes in New Zealand school children—G. M. VAUGHAN
2. The antecedents of two varieties of vocational interests—L. E. TYLER
Personality in old age—P. E. SLATER AND H. A. SCARR
Variables affecting experimenter bias in a group situation—R. ROSENTHAL, C. J. FRIEDMAN, C. A. JOHNSON, K. FODE, T. SCHILL, R. C. WHITE, AND L. L. VIKAN
Ordinal position and sex role identification—B. G. ROSENBERG AND B. SUTTON-SMITH
Personality correlates of accuracy of self-appraisal in four-year-old children—PAUL S. WEINER

VOLUME 71—January-June, 1965

1. Author index for Genetic Psychology Monographs 1926-1964—L. W. FERGUSON
A theoretical, developmental, and clinical investigation of embarrassment—J. M. SATTLER
School readiness as evaluated by Gesell developmental, visual, and projective tests—F. L. ILG, L. B. AMES, AND R. J. APPELL
Rorschach scales of regression in psychosis—S. NEIGER, A. G. SLEMON, AND D. A. QUIRK
The effects of extra tactile stimulation on a group of institutionalized infants—L. CASLER
Human-eyelid conditioning: The recent experimental literature—L. E. ROSS AND T. F. HARTMAN
2. Social status and schoolroom aggression in third-grade children—R. TOIGO
Overt and covert homosexual problems among male college students—L. J. BRAATEN AND C. D. DARLING
Role conceptions of the psychiatric aid—M. P. LAWTON AND A. E. GOLDMAN
Changes in Study of Values scores during the four years of college—W. HUNTLEY

VOLUME 72—July-December, 1965

1. A study of process and method in three-year-old twins and triplets learning to read—W. FOWLER
Individualized teaching with therapeutic aims: A methodological study—M. KOHN AND M. RUDNICK
Ability and social adjustment at midlife of persons earlier judged mentally deficient—E. L. MILLER
2. Longitudinal changes in intolerance and authoritarianism for subjects differing in amount of college education over four years—W. T. PLANT
An NDEA counseling and guidance institute: Prediction, performance, and follow-up—R. H. DANA
The influence of language on classification: A theoretical model applied to normal, retarded, and deaf children—H. G. FURTH AND N. A. MILGRAM
Time estimates as affected by incentive class and motivational level—M. FELIX
Clinical psychology: An historical survey—V. S. SEXTON

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

Statement of ownership, management and circulation required by the Act
of October 23, 1962, Section 4369, Title 39, United States Code

Date of Filing: 9/3/65

Title of Publication The Journal of Social Psychology

Frequency of issue: Published Feb., April, June, August, October, December

Location of known office of publication (Street, City, County, State, Zip Code): The Journal Press, 35 New Street, Worcester, Massachusetts 01605.

Location of the Headquarters or General Business Offices of the Publishers (Not Printers): The Journal Press, 2 Commercial Street, Provincetown, Mass. 02657.

Names and Addresses of Publisher, Editor and Managing Editor:

Publisher (Name and Address): The Journal Press, 35 New Street, Worcester, Massachusetts 01605.

Editor (Name and Address): Leonard W. Doob, 2 Commercial Street, Provincetown, Mass. 02657.

Managing Editor (Name and Address): Powell Murchison, 2 Commercial Street, Provincetown, Mass. 02657.

Owner (If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given.)

Name and Address: The Journal Press, Inc., 2 Commercial St., Provincetown, Mass.; Dorothea P. Murchison, 2 Commercial St., Provincetown, Mass.; Powell Murchison, 2 Commercial St., Provincetown, Mass.; Marjorie E. Monteith, 19 Hollins Drive, Santa Cruz, California.

Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 percent or more of Total amount of Bonds, Mortgages or other Securities (If there are none, so state): Name and Address: SAME AS ABOVE

Paragraphs 7 and 8 include, in cases where the stockholder or security holder appears upon the books of the company as trustee or in any other fiduciary relation, the name of the person or corporation for whom such trustee is acting, also the statements in the two paragraphs show the affiant's full knowledge and belief as to the circumstances and conditions under which stockholders and security holders who do not appear upon the books of the company as trustees, hold stock and securities in a capacity other than that of a bona fide owner. Names and addresses of individuals who are stockholders of a corporation which itself is a stockholder or holder of bonds, mortgages or other securities of the publishing corporation have been included in paragraphs 7 and 8 when the interests of such individuals are equivalent to 1 percent or more of the total amount of the stock or securities of the publishing corporation.

This item must be completed for all publications except those which do not carry advertising other than the publisher's own and which are named in sections 132.231, 132.232 and 132.233, postal manual (Sections 4355a, 4355b, and 4356 of Title 39, United States Code)

A. Total No. Copies printed (Net Press Run): Average No. Copies each issue during preceding 12 months 1864. Single issue nearest to Filing Date 1935.

B. Paid Circulation: 1. Sales through dealers and carriers, Street Vendors and Counter Sales: Average no. copies each issue during preceding 12 Months 0. Single issue nearest to Filing Date 0. 2. Mail Subscriptions: Average No. Copies each issue during preceding 12 Months 1668. Single issue nearest to Filing Date 1748.

C. Total Paid Circulation: Average No. Copies each issue during preceding 12 Months 1668. Single issue nearest to Filing Date 1738.

D. Free Distribution (including samples) By Mail, Carrier or other means: Average No. Copies each issue during preceding 12 Months 10. Single issue nearest to Filing Date 10.

E. Total Distribution (Sum of C and D) Average No. Copies each issue during preceding 12 Months 1668. Single issue nearest to Filing Date 1748.

F. Office use, left-over, unaccounted, spoiled after printing: Average No. Copies each issue during preceding 12 Months 196. Single issue nearest to Filing Date 187.

G. Total (Sum of E & F—should equal net press run shown in A): Average No. Copies each issue during preceding 12 Months 1864. Single issue nearest to Filing Date 1935.

I certify that the statements made by me above are correct and complete. (Signature of editor, publisher, business manager, or owner)

(Signed) Powell Murchison, Business Manager

SPECIAL ANNOUNCEMENT

The Journal Press is proud to announce that Leonard W. Doob of the Department of African Studies at Yale University has become the Assessing Editor of *The Journal of Social Psychology*. Dr. Doob is well known for his cross-cultural interests and activities, and his appointment reflects the continued interest of the management in encouraging cross-cultural research. *The Journal of Social Psychology*, as in recent years, will grant prior publication to works of this nature, and they will appear at the beginning of each issue whenever possible.

Because of the growing complexity of the field of psychology, the members of the Board of Directors have wholeheartedly agreed that several specialists, instead of one individual, should share responsibility for publishing the periodicals of The Journal Press. Active, top-flight psychologists are not likely to be available for performing the routine of copy editing. So the new position of Assessing Editor has been created. The Copy Editing Department in Provincetown will work directly with the Assessing Editor, who will be free to pursue his prime function of improving the standards for accepted work.

The process of acceptance or rejection will be a joint one. The Managing Editor will make a final decision based on two reports: one from the Assessing Editor or a member of the Editorial Board on the merits and significance of the manuscript, and one from the Copy Editing Department on the mechanical, editorial, and stylistic problems involved in publishing the paper.

Readers of *The Journal of Social Psychology* should realize that a backlog of accepted manuscripts exists and that, during a transition period, the Assessing Editor is not responsible for the contents.

\$30.00 per year
\$15.00 per volume
Single Numbers \$7.50

BIMONTHLY
Three volumes per year

December, 1965
Volume 67, Second Half

Founded in 1929 by John Dewey and Carl Murchison

THE JOURNAL OF SOCIAL PSYCHOLOGY

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

DECEMBER, 1965

(Manuscripts are printed in the order of final acceptance)

- A comparison of the direction-of-perception technique with the Likert method
in the measurement of attitudes 189
BY BARBARA L. DRINKWATER
- Discrimination of geometrical patterns in tribal, rural, and urban children . . . 197
BY P. C. NANDA, J. P. DAS, AND H. K. MISHRA
- Effect of differential valuation on group level of aspiration, decision time,
and productivity 201
BY FLORENCE DENMARK, DOROTHY MURGATROYD, AND ALBERT PEPITONE
- Bias in evaluating the presidential debates: A "Splinter" effect 211
BY RALPH L. ROSNOW
- Social-group membership, leadership, and birth order 221
BY REGINALD G. SMART

(OVER)

Copyright © 1965, by The Journal Press
Provincetown, Massachusetts, U. S. A.

All rights reserved

Published at 35 New St., Worcester, Mass.
Second-class postage paid at Worcester, Mass.

Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts

Effect of discrepant communications on self-perception	227
BY DEMETRIOS PAPAGEORGIS AND BRIAN M. McCANN	
Diagnostic use of the interaction profile	239
BY ROBERT F. BALES AND A. PAUL HARE	
Correlates of civil-rights participation	259
BY LESTER CARR AND S. OLIVER ROBERTS	
Variation in time descriptions and need achievement	269
BY ROBERT H. KNAPP AND JOHN T. GARBUTT	
A study of prejudice in Negro and white college students	273
BY LUIS PROENZA AND BONNIE R. STRICKLAND	
Studies in social attitudes: I. Variations in latitudes of acceptance and rejection as a function of varying positions on a controversial social issue	283
BY LUTFY N. DIAB	
Studies in social attitudes: II. Selectivity in mass communication media as a function of attitude-medium discrepancy	297
BY LUTFY N. DIAB	
Studies in social attitudes: III. Attitude assessment through the semantic-dif- ferential technique	303
BY LUTFY N. DIAB	
Experimental social psychology and its neighbors	315
BY EZRA STOTLAND	
Adolescent response to religious appeal as related to IPAT anxiety	325
BY C. EWING COOLEY AND JERRY B. HUTTON	
Familiarity, dogmatism, and reported student attitudes toward the disabled	329
BY JACK K. GENSLOW AND FRANK D. MAGLIONE	
Demand characteristics vs. physiological needs: A methodological critique of the "new look" school of perception	343
BY ARLENE C. BRANDON AND EDWARD E. SAMPSON	
Measurement of radicalism-conservatism	357
BY ANDREW L. COMREY AND JOHN A. NEWMAYER	
The language of postural interpretation	371
BY ROBERT H. KNAPP	
Modes of conformity and career selection of rural and urban high-school seniors	379
BY EUGENE L. GAIER AND WILLIAM F. WHITE	
BOOKS RECENTLY RECEIVED	393

A COMPARISON OF THE DIRECTION-OF-PERCEPTION TECHNIQUE WITH THE LIKERT METHOD IN THE MEASUREMENT OF ATTITUDES*

*Department of Physical Education, University of
California, Santa Barbara*

BARBARA L. DRINKWATER

A. INTRODUCTION

In a recent study relating attitudes of college teachers to teaching effectiveness two techniques of attitude measurement were employed. One, the Likert method of summated ratings, has been widely used; and considerable data on its validity and reliability are available in the literature. The other, the direction-of-perception technique, is relatively new; and information regarding its usefulness as a tool of attitude measurement is not as readily available. The use of both techniques in a single study should provide further information regarding the relative merits of the two techniques.

1. *The Likert Technique*

In the Likert method of attitude measurement the respondent is asked to check one of five possible answers to each of a series of statements. In this study the choices, preceded by their respective weights, were as follows: 1) I strongly disagree, 2) I disagree, 3) I am undecided, 4) I agree, and 5) I strongly agree. An individual's attitude score is the sum of the point values for the choices selected.

2. *The Direction-of-Perception Technique*

The direction-of-perception technique (1) is an adaptation of Hammond's error-choice method (2) in which a series of multiple-choice questions are prepared, with foils equidistant from the truth or at two opposite extremes if the true answer is indeterminate. An attitude is considered an affective state that may produce error in perception and recall, and an individual's attitude is determined by measuring the constancy of error in his choice of answers (2).

Bernberg objected to the emphasis on "error" in perception because, "To assume error or distortion in perception places a value upon the attitude which

* Received in the Editorial Office, Provincetown, Massachusetts, on January 13, 1964. Copyright, 1965, by The Journal Press.

is considered good or bad depending on the views of the test constructor or persons external to the situation" (1, p. 399). Bernberg, instead of basing his questions on fact, selected stimulus materials with no factual basis, theorizing that a respondent would choose answers in line with the direction of his perception as determined by his need-value system. An example of the type of question is found in Bernberg's survey of employee morale in a large aircraft-manufacturing plant:

When workers have worked side-by-side for a considerable length of time each individual worker becomes friendly with about: (a) 65% of the group; (b) 80% of the group (1, p. 402).

Bernberg obtained evidence for the validity of the direction-of-perception technique by comparing management ratings and union ratings of morale with the test results. The test scores corroborated the judgment of management and the union ratings and had a split-half reliability of .86.

3. *Selection of Foils*

Little information has been published on the selection of foils either for the error-choice method or the direction-of-perception method, particularly when no factual answer is available. Hammond's and Wechsler's choices of responses (2, 6) were numerically equidistant from the true answer when such answers existed or equidistant from two alternate extremes when no true answer existed. Bernberg (1) provided two types of answers: (a) numerical percentages and (b) verbal statements. After the administration of Bernberg's test, some respondents said that ". . . there was too small a difference in the percentages used in some of the questions" (1, p. 404). This comment suggested the need for inquiry into the effect on responses of the numerical separation of the foils.

B. PROCEDURE

The comparison of the Likert method of summated ratings and the direction-of-perception technique was part of a larger study on the relationship of teacher attitudes and the effectiveness of college teaching. Each of the two attitude instruments contained components designed to measure (a) attitude toward college students, (b) attitude toward the technical aspects of teaching, and (c) attitude toward student ratings of teacher effectiveness.

1. *Subjects*

The subjects were 410 teachers randomly selected to represent universities, liberal-arts colleges, and teachers colleges of the midwestern states.

2. Selection of Attitude Items

Current literature pertaining directly to college students was used as the basis for preparing an Attitude Inventory of 36 statements to measure teacher attitudes toward college students. In general the statements sample opinion on student behavior, ideals, attitudes, academic interest, and the nonacademic responsibilities of the teacher to his students. Effort was made to phrase the statements in such a way that the direction of the favorable response was obscured. Corresponding questions were prepared for the Questionnaire for College Teachers, the direction-of-perception instrument. The order of the items on both instruments was the same and was determined by a random selection of the index cards on which the statements had been typed.

Twenty-four items to measure attitude toward the technical aspects of teaching were selected from studies listing the instructional competencies relevant to classroom teaching procedures: (a) preparation for class, (b) testing procedures, (c) conduct of the class, (d) the use of teaching aids, (e) rapport with students, and (f) educational methods.

Because the cooperation of the respondents depended in part on their willingness to be evaluated by students, a section was included dealing with the respondent's attitude toward student ratings of teacher effectiveness. The statements used were taken from studies reporting results of research on many of the objections to student ratings and from the *Manual of Instruction for the Purdue Rating Scale for Instruction* (4). The results from this portion of the attitude instruments were used to check the bias to be expected if only teachers possessing favorable attitudes toward student ratings participated.

3. Selection of Direction-of-Perception Foils

Hammond's suggestion of using more than two answers for each question (2) was incorporated to provide some indication of attitude intensity. Four methods were used to determine the numerical separation of the answers. On 30 questions four answers were separated by a single value ranging from seven to 21 percentage points. For example in the following question, each answer is 21 percentage points distant from adjacent answers:

What per cent of college students indicate that they would take advantage of a teacher's absence from the room to cheat on an examination? (a) 21%, (b) 42%, (c) 63%, (d) 84%.

Fifteen other questions had foils separated by a single value but were arranged so that corresponding favorable and unfavorable choices were equidistant from a value of 50 per cent.

What per cent of college teachers would prefer to have their effectiveness judged by their department heads rather than by their students?

(a) 82%, (b) 66%, (c) 34%, (d) 18%.

On 12 questions answers were not separated by any common value, but were distributed so that those answers equally favorable or unfavorable were equidistant from a value of 50 per cent.

What proportion of college students condone cheating as an acceptable practice? (a) 33%, (b) 45%, (c) 55%, (d) 67%.

On 14 questions the favorable choices were separated by four points, the distance between a favorable and unfavorable choice was six points, and unfavorable choices were four points apart.

What proportion of college teachers feel that the ability to speak "off-the-cuff" is more important than detailed preparation for class? (a) 25%, (b) 29%, (c) 35%, (d) 39%.

Ten questions had miscellaneous types of answers in which, for example, the choices were expressed in terms of "greater," "higher," "more," or "always."

The variation in the number of questions in each category resulted from the necessity of providing the questionnaire with face validity. For example, in answer to some of the questions answers as large as 50 per cent would have appeared unrealistic. To determine whether the methods of selecting foils affected the responses, the answers within each category were compared with replies to the corresponding Likert items.

4. *Comparability of the Attitude Instruments*

To determine if the corresponding items on the two instruments were comparable in content, a combined form of the two inventories (with questions from the direction-of-perception instrument in Column A and corresponding Likert statements in Column B) was presented to a class of 30 graduate students. The students were asked to read both the question and the statement carefully and to indicate on a five-point scale how well they felt the two agreed in content. The scale choices were as follows: (a) no relationship between attitude statement and question, (b) very little relationship between attitude statement and question, (c) a fair degree of relationship between the attitude statement and question, (d) attitude statement and question cover approximately the same topic, and (e) attitude statement and question cover precisely the same topic. Sixty-two of the paired items had an average score of 4.00 or better, while the remaining 19 scored between 3.50 and 3.99. Because all the matched items were judged to have better than

a fair degree of relationship, it was decided to retain all 81 items on each attitude instrument.

C. RESULTS

The reliability coefficients of the attitude instruments shown in Table 1 are coefficients of equivalence as estimated by the split-half method and corrected by the Spearman-Brown formula. With the exception of Area 2,

TABLE 1
RELIABILITY COEFFICIENTS FOR THE ATTITUDE INVENTORY
AND THE QUESTIONNAIRE FOR COLLEGE TEACHERS

Scores	Attitude Inventory		Questionnaire	
	Split-halves <i>r</i>	Corrected <i>r</i>	Split-halves <i>r</i>	Corrected <i>r</i>
Area 1	.70	.82	.77	.87
Area 2	.38	.55	.47	.64
Area 3	.81	.90	.72	.84
TOTAL	.75	.85	.83	.91

attitude toward the technical aspects of teaching, the reliability coefficients of the subscales in both instruments indicate that they may be used independently of the total scores as measures of the attitudes of college teachers toward their students and toward student ratings of teacher performance.

A comparison was made of the reliability coefficients for the Questionnaire for College Teachers (QCT) which use the direction-of-perception technique, and the Attitude Inventory (AI), an example of the Likert method of attitude measurement. The results are shown in Table 2. Based on total scores, the reliability of the Questionnaire is significantly higher than that of the Inventory at the .05 level. There is no significant difference in the reliabilities for Area 1. The reliability for Area 2, though comparatively low, is higher (at the .05 level) on the Questionnaire. For Area 3, the reverse is true; the reliability of the Inventory subscale is significantly higher (at the .05 level) than that of the Questionnaire.

TABLE 2
COMPARISON OF RELIABILITY COEFFICIENTS FOR QUESTIONNAIRE
FOR COLLEGE TEACHERS AND THE ATTITUDE INVENTORY
USING FISHER'S *z* FUNCTION

	Questionnaire	Inventory	Difference (<i>S.E._D</i> = .11)
Area 1	1.33	1.16	.17
Area 2	.76	.50	.26*
Area 3	1.22	1.47	-.25*
TOTAL	1.53	1.26	.27*

* Significant at .05 level.

The results of the correlation of the total and area scores on the Questionnaire for College Teachers with those of the Attitude Inventory are shown in Table 3. An r of .75 between total scores makes it possible to reject (at the

TABLE 3
CORRELATION OF TOTAL AND AREA SCORES OF THE QUESTIONNAIRE
FOR COLLEGE TEACHERS WITH CORRESPONDING
SCORES ON THE ATTITUDE INVENTORY

	Area 1	Area 2	Area 3	TOTAL
r	.78**	.31*	.67**	.75**

* Significant at .05 level.

** Significant at .01 level.

.01 level) the hypothesis that the correlation between the Likert-type attitude scale (AI) and the direction-of-perception attitude inventory (QCT) is zero. Significant correlations between the scores of subscales also make it necessary to reject the hypothesis that the correlations between subareas on the Likert-type attitude scale (AI) and the corresponding subareas on the direction-of-perception scale are zero. For Areas 1 and 3 this hypothesis is rejected at the .01 level; for Area 2, at the 5.0 level. From these results it appears that there is a definite relationship between the Likert method of attitude measurement and the direction-of-perception technique.

Five methods were used in preparing the responses to the items on the Questionnaire for College Teachers. From the results reported in Table 4 it is evident that each of the five methods produces results comparable to those obtained for corresponding items on the Likert-type Attitude Inventory. An

TABLE 4
CORRELATION OF ITEMS ON QCT GROUPED ACCORDING TO METHOD USED
IN FORMATION OF THEIR RESPONSES WITH CORRESPONDING
ITEMS IN THE ATTITUDE INVENTORY

	Methods				
	a	b	c	d	e
r	.67**	.59**	.58**	.34**	.59**

** Significant at .01 level.

analysis of the differences between these r 's, as shown in Table 5, results in only one significant difference—that between Method a (foils separated by a single value) and Method d (foils separated by 4-6-4 percentage points). Before Method d is discarded as a means of formulating direction-of-perception foils, it should be considered that most items of this type are in Area 2 and that the lower reliability of this area may have had an adverse effect on the comparison of these items with corresponding Likert items. No difference between the remaining four methods was significant (at the .05 level).

TABLE 5
DIFFERENCES OF "METHODS" r EXPRESSED IN FISHER'S z

	a	b	Methods c	d	e	$S.E._D = .17$
a		.13	.15	.46**	.13	
b			.02	.33	.00	
c				.31	-.02	
d					-.33	

** Significant at .01 level.

One of the difficulties mentioned in relation to the direction-of-perception method was the difficulty in formulating items (6). In many cases data are not available to serve as a basis for the preparation of responses that are equidistant from the truth. Since the items in this study, none of which were based on actual data, were significantly related to the corresponding Likert-type statements, it is evident that lack of data need not prohibit the use of the direction-of-perception technique.

One other point to be considered is that both Hammond (2) and Wechsler (6) combined factual and nonfactual items on their attitude instruments. This procedure resulted in comparatively long tests, but in a small number of attitude statements. Hammond (2), for example, suggested a ratio of two factual questions to each nonfactual or attitude item. As a result of the small number of attitude items, the reliability coefficients for these instruments were relatively low. No factual items were included in the Questionnaire for College Teachers, and all the reliability coefficients (with the exception of Area 2) were higher than those reported by Hammond (2) or Wechsler (6). It appears that a combination of factual and nonfactual items is not necessary in the construction of an attitude instrument using the direction-of-perception method.

D. SUMMARY

Two attitude inventories, one using the Likert method and the other the direction-of-perception technique, were developed to measure the attitudes of college teachers toward college students, the technical aspects of teaching, and student ratings of teacher effectiveness. The reliabilities of these instruments were determined from responses obtained from college and university teachers in the midwestern states. Eighty-five teachers returned the Questionnaire for College Teachers (an example of the direction-of-perception technique) and 74 returned the Likert-type Attitude Inventory. The reliability coefficients for Areas 1 and 3 and for the complete instruments were satisfactory. A lower reliability for Area 2 suggests that caution be used in basing any conclusions upon data derived from it.

The data obtained from the administration of these three instruments were analyzed as follows: (a) the relationship of the Attitude Inventory and the Questionnaire for College Teachers was studied in terms of the total scores and the scores for each attitude area and the reliability coefficients; (b) scores grouped according to the methods used in preparing the answers to the direction-of-perception questions were compared with scores of corresponding items on the Likert-type scale.

For the population defined in this study the data support the following conclusions:

1. The Questionnaire for College Teachers and the Attitude Inventory are reliable measures of attitudes toward college students and toward student ratings of teacher performance. The area on both instruments designed to measure attitudes toward the technical aspects of teaching should be further refined before data derived from its administration can be used with assurance.

2. Attitude scores as measured by the direction-of-perception technique are significantly correlated with attitude scores measured by the Likert method. These results, plus the adequate reliabilities on two of the three subscales suggest that the direction-of-perception technique is a reliable method of measuring attitudes.

3. All of the five methods used in formulating responses to the questions on the Questionnaire for College Teachers produced responses comparable to those obtained on the Attitude Inventory. It is not necessary to have factual data as a basis for preparing foils or to combine factual and nonfactual items in order to have a reliable instrument.

REFERENCES

1. BERNBERG, R. E. The direction of perception technique of attitude measurement. *Int. J. Opin. & Attit. Res.*, 1951, **5**, 397-406.
2. HAMMOND, K. R. Measuring attitudes by error-choice: An indirect method. *J. Abn. & Soc. Psychol.*, 1948, **43**, 38-48.
3. LIKERT, R. Technique for the measurement of attitudes. *Arch. Psychol.*, 1932, **22**, 5-43.
4. REMMERS, H. H. Manual of Instructions for the Purdue Rating Scale for Instruction. West Lafayette, Indiana: Univ. Book Store, 1960.
5. WANG, C. K. A. Suggested criteria for writing attitude statements. *J. Soc. Psychol.*, 1932, **3**, 367-373.
6. WECHSLER, I. R. An investigation of attitudes toward labor and management by means of the error-choice method. *J. Soc. Psychol.*, 1950, **32**, 51-62.

Department of Physical Education
University of California, Santa Barbara
Santa Barbara, California

DISCRIMINATION OF GEOMETRICAL PATTERNS IN TRIBAL, RURAL, AND URBAN CHILDREN*

Utkal University, Bhubaneswar, India

P. C. NANDA, J. P. DAS,¹ AND H. K. MISHRA

A. INTRODUCTION

Does past experience with a certain pattern alone always lead to its correct discrimination from another pattern closely resembling the original? Is this true even if distinct verbal labels are not available for the two patterns? In the present study, three samples of children from three separate cultural environments were tested for their ability to discriminate between the following pairs of figures: circle *vs.* ellipse, rectangle *vs.* parallelogram, obtuse-angled *vs.* acute-angled triangle, and a circle within a triangle *vs.* a triangle within a circle. The children ranged in age between 6 and 10 and lacked appropriate verbal labels for the test figures.

The study was conducted in three parts, starting with the children from an aboriginal tribe in the southern part of Orissa, India. In the summer of 1962, the three authors went together to the interior of a tribal area. As their entry into a tribal village attracted adults and children alike, one of the authors kept the villagers amused with view-master pictures while the other two authors tested subjects in groups of five. In the remote villages visited, any group of strangers could be perceived by the children as official vaccinators, therefore, the children's fears had to be allayed by a generous distribution of sweets. Rapport with the parents of the tribal children was easily established by the authors introducing themselves as teachers from a "big school." Rapport was essential, so that shy children would be encouraged by their parents to participate in the experiment. No particular problem was encountered in obtaining the rural and urban samples.

B. METHOD

1. *Experimental Design*

Each of the three samples was divided into eight experimental groups, following the requirements of a two-by-two-by-two factorial design in a man-

* Received in the Editorial Office, Provincetown, Massachusetts, on March 4, 1964. Publication delayed for technical reasons. Copyright, 1965, by The Journal Press.

¹ This manuscript was prepared during the tenure of the author as a Kennedy Foundation Professor (1963-1964) at George Peabody College for Teachers.

ner to be indicated. The four pairs of geometrical figures were classified under two broad categories as follows: (a) parallelogram *vs.* rectangle and ellipse *vs.* circle and (b) obtuse-angled *vs.* acute-angled triangles and circle within triangle *vs.* triangle within circle. The figures in Category (a) were presented to a group of Ss half of whom had the parallelogram and ellipse as cues associated with reinforcement, whereas the other half had rectangle and circle. A similar design was followed for the stimuli in the second category. Further, the authors made an attempt to determine whether or not filled figures (inside area blackened by India ink) were more easily discriminated than unfilled figures (contours drawn with India ink on a white background). Thus, each of the categories was presented in two versions, so that eight experimental groups were required. Five children in each sample were randomly assigned to each of the eight cells.

2. Subjects

The tribal children were from two tribal villages. The villages can be called remote because they are not approachable by roads and do not have an elementary school or hospital within a distance of five miles. The children were illiterate and could never have come across geometrical pattern in print. Children between the ages of 6 and 10 were to be chosen, but a correct statement of age could not always be obtained from parents. Therefore, the authors selected as Ss children who had cut at least one new tooth (approximate age about 6 years) but who had not cut all of their new teeth (about 10 years).

The illiterate-village sample was selected from a remote nontribal village where most children do not go to school and the percentage of illiteracy in the adult population is high. The authors would have liked to test urban-illiterate children comparable in socioeconomic status to the tribal and rural children. But school going is so common in the towns that the authors could not find sufficient children to qualify for the desired sample.

3. Procedure

As indicated earlier, subjects were tested in groups of five with two pairs of stimuli in two experimental sessions 30 minutes apart. In one of the sessions, Ss were required to discriminate between the elements of a pair of stimuli (e.g., parallelogram *vs.* rectangle) in a continuous session during which the pair was presented for a maximum number of 16 trials or until S reached a criterion of five consecutively correct responses. A stimulus pair for discrimination was presented on a tray with slots, so that S could be

instructed to pull out the figure that he thought concealed a sweet. *S* was allowed to collect the sweet and save it until the experiment was over. As mentioned earlier, the groups were counterbalanced on the reinforced figure, so that half of the *Ss* received a sweet under, say, parallelogram; while the other, under a rectangle. *Ss* who had attempted to discriminate between a parallelogram and a rectangle were not asked until 30 minutes later to begin their second task: i.e., discriminating between a circle and an ellipse.

C. RESULTS AND DISCUSSION

S's responses were scored in two ways: (a) failure to discriminate within 16 trials and (b) the number of trials required to reach the criterion of five consecutive correct responses. Failure to discriminate was assigned a score of 16.

To prepare a table suitable for chi-square analysis, the two conditions of reinforcement were neglected; otherwise the expected frequency per cell would have been too small. Further, the results of reinforcing the circle, for example, do not seem to differ from those when the ellipse was reinforced. The total number of failures for the urban, rural, and tribal samples were 12, 34, and 62 respectively [see Table 1]. With two degrees of freedom, chi

TABLE 1
NUMBER OF FAILURES*

Type of figure	Urban		Rural		Tribal	
	(a)	(b)	(a)	(b)	(a)	(b)
Filled	3	4	11	6	12	15
Unfilled	1	4	10	7	20	16

* See text, p. 198, for definitions of Categories (a) and (b).

square (34.89) is significant beyond the .01 level; therefore, ability to discriminate geometrical patterns varies directly with the degree of urbanization.

In a three-by-two-by-two analysis of variance without replication [see (1)], the reinforcement-treatment level was ignored because of a large number of failure scores. To avoid undue strain on the assumptions of analysis variance individual scores were pooled for each cell, and the triple interaction between samples (filled *vs.* unfilled figures) and Categories (a) and (b) was taken as the error term. No significant *F* was obtained [see Table 2].

The above results suggest that a more sensitive measure, like the number of trials required to discriminate the figures, fails to show any difference in the performance of the three cultural groups. But the distribution of the total number of *Ss* failing to discriminate suggests geometrical patterns are more

TABLE 2
SUMMARY OF ANALYSIS OF VARIANCE

Source	<i>df</i>	<i>MS</i>	<i>F</i> *
Samples	2	36457.	7.083
Type†	1	1200.	
Shape‡	1	363.	
Interaction			
Samples-by-type	2	1027.	
Samples-by-shape	2	2428.	
Type-by-shape	1	33.3	
Samples-by-type-by-shape	2	5147.	
Total	11	86825.	

* Except for the one value given (7.083), no *F* is larger than one.

† Filled figure *vs.* unfilled figure.

‡ Category (a) *vs.* Category (b). See text, p. 198.

efficiently perceived by children from "civilized" cultures. The level and not the speed of learning turns out to be a reliable measure for use in the crosscultural samples.

REFERENCE

1. EDWARDS, A. L. *Experimental Design in Psychological Research*. New York: Holt, 1960.

Department of Psychology

Utkal University

Bhubaneswar, India

EFFECT OF DIFFERENTIAL VALUATION ON GROUP LEVEL OF ASPIRATION, DECISION TIME, AND PRODUCTIVITY*

Queens College, Wagner College, and University of Pennsylvania

FLORENCE DENMARK, DOROTHY MURGATROYD, AND ALBERT PEPITONE

A. INTRODUCTION

Two aspects of human-aspiration behavior are (*a*) the level at which goals are set by individuals and (*b*) the persistence that individuals show in setting goals for themselves. Both kinds of behavior have been shown to be sensitive to social influence: for example, the level at which task goals are set is strongly affected by information given to the individual concerning the performance in that task by groups of known status levels [Chapman and Volkman (1)]. The persistence of aspiration, in terms of how many trials the individual continues to set task goals, is affected by the direction and magnitude of the discrepancy between his performance and that of a peer group [Dreyer (2)]. In both instances, the individuals can be said to conform to the social influence; however, the latter may be considered entirely passive because there has been no attempt to exert influence.

Looking at aspiration behavior as a case of social conformity, a number of questions arise analogous to questions of opinion and attitude conformity: for example, can aspiration behavior also be affected by active social influence? Then, too, active social influence presumably could occur on different levels: for example, direct influence refers to the instance in which one individual persuades another to set a goal at a given level. An indirect level of active influence might occur by way of an evaluation of the individual's ability. The focus of the present study is on this more subtle social determinant of aspiration behavior.

Our working hypothesis was that the level of difficulty of a task goal to which an individual aspires, in the absence of any experience with the task, is a function of the individual's relevant ability, as evaluated by a competent authority. The second question was whether or not such influence represents change in self-evaluation. Clearly, a shift (or difference between conditions) in aspiration could be a response that accommodates the source of influence only so long as the latter is psychologically present, can mete out punishment

* Received in the Editorial Office, Provincetown, Massachusetts, on April 17, 1964. Copyright, 1965, by The Journal Press.

for noncompliance, or keeps the individual dependent upon him for evaluation.

Theoretically, a person's self-evaluation depends upon his ability, as reflected by achievement; however, such measurement presupposes either an objective score on the task or the opportunity for comparison with other individuals. When such standards are lacking, achievement (hence, the individual's ability) is not interpretable. To the extent that achievement approaches an authoritative evaluation or if achievement cannot be evaluated, one would predict acceptance of the influence; that is, self-evaluation should parallel aspiration behavior, and both would be a function of the external evaluation of the individual.

B. HYPOTHESES

1. When informed that they have high ability with regard to a particular task, groups who lack knowledge about their ability for that task choose higher levels of difficulty at which to perform the task than do groups who have been told by a competent authority that their ability is average or low.

2. When informed that they have average ability with regard to a particular task, groups who lack knowledge about their ability for that task choose higher levels of difficulty at which to perform the task than do groups who have been told by a competent authority that their ability is low.

3. When informed that they have high ability with regard to a particular task, groups who lack knowledge about their ability for that task rate their own ability to perform the task (self-evaluation) as greater than do groups who have been told by a competent authority that their ability is average or low.

4. When informed that they have average ability with regard to a particular task, groups who lack knowledge about their ability for that task rate their own ability to perform the task (self-evaluation) as greater than do groups who have been told by a competent authority that their ability is low.

C. METHOD

1. *Subjects*

The subjects (Ss) in this study were 60 pairs of college women enrolled in general psychology at the University of Pennsylvania. By participating, the students were excused from the preparation of one laboratory report.

2. *Procedure*

Ss filled out a personal-data questionnaire and took a timed spatial-relations test during the recruitment session. They were informed that the experimenter

was concerned with how people reach decisions regarding a task and how they work together at it. The personal-data sheet requested information on items such as average grade in high school and college, major subject in college, extracurricular activities, the kind of person *S* was, and the kind of person *S* would like to work with. *Ss* were paired at random without reference to any of the preceding information.

a. Experimental sessions. Each pair of *Ss* met by appointment with *E* to work on a card-sorting task. At that time, *Ss* were informed that the particular task involved had proved useful in previous studies investigating the way in which people reach decisions and work together. The task was that of sorting 35 specially designed cards, permanently mounted on a large cardboard. Each card could be classified according to four characteristics: the size of the card, the size and density of the dots covering the surface of the card, the number of straight lines forming an asterisk-like figure on each card, and the pattern of points where each of three heavy black lines reached the edges of each card. *Ss* were to sort each of these characteristics into as many different categories as they decided were necessary, working on one characteristic at a time.

Next, *Ss* were presented with a typed sheet containing presumed levels of difficulty for performing the task. They were informed that an earlier study had established certain orders of working on the task as more difficult than others. For instance, an order that began with categorizing the size of the cards was more difficult than one that began with categorizing asterisks. Only one sheet was given to the pair of *Ss*, so that interaction would be induced by the necessity of the *Ss* working together. Nine levels of difficulty were typed on the sheet, with No. 1 designated as the most difficult and No. 9 as the least difficult. The other levels supposedly were scaled between the extremes. To emphasize the presumed scaling of difficulty and to give a further basis for discussion, two other scales were presented along with the nine levels. *E* then told *Ss* that, upon comparing the information on their personal-data sheets and their scores on the spatial-relations test previously administered with those of others who had already performed the task, they were judged to have high, average, or low ability.

Next, *Ss* were instructed that they had to reach a joint decision about the order or level at which they wished to work, taking into consideration the various risks involved. *E* left the room while *Ss* reached their decision, and the time required was recorded by an observer. When the decision was reached, *E* returned and started *Ss* on the task. *Ss* worked together, telling *E* by number what cards they wished placed in each of the categories they had

decided upon. *E* kept a record of the categorized cards and made various observations during the decision making and the work session. The time limit for the task was 20 minutes.

At the end of 20 minutes, a questionnaire was administered to determine to what degree the experimental induction had been effective. Questions concerned self-estimates of ability regarding the task, the amount of attraction to the group, and whether or not prior acquaintance with the partner existed.

D. RESULTS

Table 1 shows the level of difficulty, decision time, and productivity for each of the treatment groups. Analyses of variance yield no *F*s significant at

TABLE 1
EFFECTS OF ABILITY INDUCTION ON LEVEL OF DIFFICULTY, DECISION TIME,
AND PRODUCTIVITY

Variable	High (<i>N</i> = 20)	Treatment Medium (<i>N</i> = 20)	Low (<i>N</i> = 20)	<i>F</i> *
Level of difficulty (mean level chosen)	2.1	4.3	3.9	9.17
Decision time (in seconds)	54.2	88.1	80.9	3.03
Gross productivity (number of categories sorted)	88.5	84.7	92.4	.50
Net productivity (gross productivity minus errors)	70.7	71.9	75.4	.42
Errors (number incorrectly sorted)	12.8	12.8	17.0	.56

* Evaluation of the significance of the overall variation among conditions is based on an *F* test with two and 57 degrees of freedom. Comparisons between treatments are based on two-tailed *t* tests. For level of difficulty, the overall test is significant at the .01 level. The difference between the high-ability group and the medium-ability group is significant at the .01 level, and the difference between the high-ability group and the low-ability group is significant at the .02 level. No other difference is significant.

the .01 level for gross productivity, net productivity, or errors. A significant *F* appears for the choice of a level of difficulty. Individual *t* tests give significant differences between the high-ability and average-ability treatment groups as well as between the high-ability and low-ability treatment groups. This same pattern tended to occur for decision time, although the difference was not statistically significant.

Table 2 gives the mean scale value for each group on the final questionnaire responses and it confirms the experimental induction in regard to ability as having been effectively perceived by the *S*s ($p < .01$). No significant differences were found between the three treatment groups for liking the other partner and degree of previous acquaintanceship, because the intention was to control these factors. An important result is the valuation the subjects put on their

TABLE 2
MEAN RESPONSE SCORES ON FINAL QUESTIONNAIRE

Question	High (<i>N</i> = 40)	Treatment Medium (<i>N</i> = 40)	Low (<i>N</i> = 40)	<i>F</i> *
1. Like to work with same partner.	1.9	2.4	1.6	2.16
2. Like partner in social situation.	2.8	2.6	2.1	2.10
3. Self-estimate.	3.1	3.6	3.6	2.38
4. Experimenter's judgment.	2.1	3.8	5.0	9.67
5. How well did you know partner?	5.2	5.0	4.7	2.00

* Evaluation of the significance of overall variation between conditions is based on an *F* test with two and 117 degrees of freedom. Comparisons between treatments are based on two-tailed tests. For experimenter's judgment, the overall test is significant at the .01 level; so is the difference between the high-ability group and the medium-ability group, the difference between the high-ability group and the low-ability group, and the difference between the medium-ability group and the low-ability group. No other difference is significant.

own ability. No significant differences appeared between the three groups, indicating that contrary to the verbal induction of *E*, *Ss* themselves took a medial position in regard to their ability.

Rank-difference correlations over all treatment groups between level of difficulty and productivity, level of difficulty and self-evaluation, level of difficulty and decision time, and self-evaluation and productivity were examined and reveal a significant rho of .39 between level of difficulty and decision time only. Differences in variability between the three treatment groups for level of difficulty, decision time, gross and net productivity, and errors were also tested. The greatest variability occurred for level of difficulty, but this *F* did not reach the .05 level of significance.

E. DISCUSSION

Table 1 confirms Hypothesis 1. *E*'s arbitrary valuation was effective in influencing the original decision of the high-ability and medium-ability groups in regard to their level of difficulty.

That the mean of the low-ability group was slightly higher than that for the medium-ability group fails to confirm Hypothesis 2 and suggests that such influence may be asymmetrical. *E* can influence aspiration upward but not downward. Such seems to have been the case in Dreyer 1954 study (2).

According to level-of-aspiration theory, in the initial choice situation the strength of the force to choose high levels of difficulty can be assumed to be dependent on the positive incentive value of success: i.e., the objective or material rewards of success plus subjective feelings of accomplishment if successful. Opposing this is the strength of the force to avoid difficult levels,

assumed to be dependent on the negative incentive value of failure: i.e., fear of failure, plus the extra effort involved in more difficult tasks. Both forces may be modified by subjective probabilities of success and failure, which might be again assumed to be some function of perceived task difficulty-ability, as noted by Lewin (5). However, with no previous knowledge of task difficulty or ability on the part of the individual or group, the 50-50 probability of success and failure then corresponds to the most-probable achievement (6). Thus, choice of a level of difficulty should be made in the middle range of a series of choices where the probabilities of success and failure are 50-50 under these conditions.

Escalona's Resultant Valence Theory (6) takes into account the greater positive valence of success in this culture, as does Festinger who, in his social-comparison processes (3), postulates a unidirectional drive to evaluate abilities upward in Western culture. Such a drive would act not only as a force for *S* to choose more difficult levels but also as a restraining force against *S*'s choosing easier levels. In the presence of this additional factor, then, the resultant choice of a level of difficulty would be above the point at which the probabilities of success and failure are 50-50.

In the present study, the experimental induction of high ability had the effect of reducing the force on *S*s to avoid difficult levels because it implied that the probability of failure at those levels was minimal. The principal force remaining was upward and decision was made accordingly.

The medium-ability induction appeared to have minimal effect on the forces already operative in the situation. Since the medium-ability groups chose a level slightly above the middle range of difficulty, the drive to evaluate abilities upward apparently influenced their decisions to this degree. For the low-ability *S*s, the results of the induction were not as expected. *S*s in this treatment group chose the highest levels of difficulty as well as those in the middle range. Since choice of high levels of difficulty was not true of medium-ability *S*s, something more than the drive to evaluate abilities upward appears operative. The unfavorable evaluation may have created a state of dissonance for the individual, as formulated by Festinger (4), and this choice behavior may have been an attempt to reduce dissonance. The tendency for anxious persons to set either extremely high or very low aspiration levels when easy success or no chance for success minimizes anxiety about failure was noted in early studies (6). Those low-ability *S*s who chose the highest levels appeared to behave in a similar fashion. When such *S*s were questioned postexperimentally, it appeared that their choice behavior was defensive in character. For instance, they almost always failed to mention the induction as influencing

their choice of level of difficulty; whereas Ss making other choices almost always did so. Other Ss in the low-ability groups appeared to have taken the experimental induction into account to some degree, but no group chose one of the lowest levels, a "floor effect" apparently stemming from the drive to evaluate abilities upward. Such Ss, however, would be expected to experience increased tension, and this expectation is borne out by the significant correlation between decision time and level of difficulty. In the case of the high-ability group, the induction facilitated their decision; while for the unrealistic low-ability Ss refusal to take the evaluation into account also resulted in a quicker decision time.

In regard to gross and net productivity and errors, the productivity of the three groups cannot be expected to be influenced by their aspiration if the task is above a critical level of difficulty or below a certain level. Clearly, if the objective constraints are strong, differences in motivation are irrelevant. By the same token, if the task is simple, variations in drive will not show up. It is only within a middle range of difficulty that aspiration differences would be permitted to be reflected in productivity scores. That the task was neither too difficult nor too simple is shown by the wide variation in the scores achieved by the Ss. However, from Table 1 it is apparent there is no overall difference in productivity across groups. There is a slight decline from high ability to medium ability, and a rise of almost eight points from medium ability to low ability. The latter again suggests the possibility of asymmetrical influence, which in this case is revealed by an extra, compensatory push for the Ss who are evaluated low.

Concerning errors, there is also no significant difference among experimental conditions. The sharp increase in errors on the part of the low-ability Ss again suggests the presence of a compulsive indiscriminate surge.

In regard to Hypotheses 3 and 4 concerning the Ss' self-evaluation (i.e., the question of the actual acceptance of influence), Question 3 in Table 2 reveals that, contrary to the hypotheses, there is no overall difference across conditions. Since the questionnaire was administered after the task had been attempted, Ss' experience with the task may have affected their ratings or the influence may have been temporary. This former possibility was tested, but no significant correlation was found between self-evaluation and productivity. The possibility that self-evaluation may have influenced level of aspiration was also examined; but here, too, no significant correlation between self-evaluation and level of difficulty was found. It appears that the Ss' self-evaluation did not depend upon actual productivity for the task nor was their self-evaluation related to level of difficulty. Verbal comments showed Ss at all levels felt

little pressure to work for a high score under the motivational conditions in this experiment.

F. SUMMARY AND CONCLUSIONS

The effect of arbitrary social evaluation as to ability on group level of aspiration was studied with 60 college women, who met in pairs with *E* to work on a card-sorting task. Three different levels of ability—high, medium, and low—were attributed to the *Ss* ostensibly on the basis of a spatial-relations test and personal-data sheet before *Ss* jointly chose a level of difficulty at which to perform the task.

Choice of level of difficulty differed significantly only for the high-ability and medium-ability treatments. The level chosen by the low-ability treatment groups was slightly higher than that chosen by the medium-ability groups, but the difference was not statistically significant. No group chose a level at or near the bottom of the scale, but there was a tendency for the low-ability treatment groups to be more variable in their choice than either of the other two groups. This tendency did not reach statistical significance, but was indicative of unrealistic goal setting under the influence of this induction.

The possibilities that level of aspiration may have been influenced by self-evaluation rather than the experimental induction or that actual performance on the task may have influenced self-evaluation were checked, but no significant correlations were found. A significant correlation between the difficulty level selected and the amount of time taken for decision appeared over all groups. Those *Ss* who chose higher levels of difficulty required less time for decision, regardless of the induction. No differences appeared between the three treatments for gross or net productivity, number of errors, or attraction to the group.

From this study it may be concluded that social influence affects group decision in regard to level of difficulty in line with its evaluation (even though arbitrary) when no other information in regard to ability is available only if the evaluation is not counter to the cultural drive to evaluate abilities upward. However, such influence appears noneffective in changing productivity at any level when motivation is otherwise minimal.

REFERENCES

1. CHAPMAN, D. W., & VOLKMAN, J. A. A social determinant of the level of aspiration. *J. Abn. & Soc. Psychol.*, 1939, **34**, 225-238.
2. DREYER, A. S. Aspiration behavior as influenced by expectation and group comparison. *Hum. Relat.*, 1954, **7**, 175-190.
3. FESTINGER, L. A theory of social comparison processes. *Hum. Relat.*, 1954, **7**, 117-140.

4. ———. A Theory of Cognitive Dissonance. Evanston, Ill.: Row, Peterson, 1957.
5. LEWIN, K. Frontiers in group dynamics, part I. *Hum. Relat.*, 1947, 1, 5-130.
6. LEWIN, K., DEMBO, T., FESTINGER, L., & SEARS, P. S. Level of aspiration. In J. McV. Hunt (Ed.), *Personality and the Behavior Disorders (Vol. 1)*. New York: Ronald, 1944.

*Department of Psychology
Hunter College in the Bronx of the
City University of New York
Bedford Park Boulevard West
New York 68, New York*

BIAS IN EVALUATING THE PRESIDENTIAL DEBATES:
A "SPLINTER" EFFECT*¹

Communication Research Division, Boston University

RALPH L. ROSNOW

A. PROBLEM

In the time since we have become aware of the possibility that the 1960 presidential debates may have swung victory to one side and away from the other, many investigators have implied the need for a full consideration of pertinent variables if the various psychological effects are to be teased out of what has popularly become the ritual of political debating. In the present paper, we shall be concerned with a problem in evaluating future presidential debates, more specifically with the role of two variables in biasing evaluation.

A multiplicity of variables conceivably might complicate a design for such an evaluation and, though much research has already been carried out, the impact on the judgmental processes of an observer of such things as his personal commitment to the stimuli he is evaluating, the order of presentation of stimuli to be evaluated, the observer's prior familiarity or unfamiliarity with the stimuli, problems in pretesting's dispositionally sensitizing observers to one or the other stimulus, and a host of other factors have not been worked out in detail with regard to the issue at hand.

The present paper deals with one of these problems—the effect of a person's "commitment" to a political party or to a political candidate on his judgments of political candidates. One question to be answered is whether politically committed observers, if so instructed, may be expected dispassionately to evaluate their own and the opposition's candidate on any given criterion of effectiveness. If it can be demonstrated that they are not so able, then any evaluative design must have built into it a component to ameliorate such error.

* Received in the Editorial Office, Provincetown, Massachusetts, on April 27, 1964. Copyright, 1965, by The Journal Press.

¹ This study was carried out while the author was with the Bureau of Social Science Research. Data analysis and write-up were carried out with support from the Communication Research Center, Boston University. A number of people contributed their time and efforts in one or more aspects of the research. Ivor Wayne was particularly helpful in all phases. Gordon W. Russell and Peter Mayer aided in both obtaining and running Ss. and Douglas Inglish of the Republican National Committee was kind enough to solicit the aid of his colleagues to be Ss. Fredric A. Powell, Edward J. Robinson, Bernard Rubin, and Gerhart D. Wiebe made valuable criticisms of an earlier version of this paper and suggested several interesting hypotheses.

In another context, the problem of personal commitment biasing personal judgmental processes has been experimentally investigated by Bass and Duntzman (1). They point out that examples of unrealistic overappraisal or underestimation of nations and of political situations are strewn throughout history. They note the United States' overevaluation of the worth of its collaborators in the Bay of Pigs incident, this country's underestimation of the Japanese during World War II, and the overevaluation in 1939 by England and France of each other but the underevaluation of their common enemy Germany. By experimentally producing a laboratory situation of conflict between allied and competing groups, Bass and Duntzman were able to demonstrate empirically a consistent underestimation by *Ss* of their competitors but a generally incremental valuation of allied groups. Results were consistent with those obtained in a similar study by Blake and Mouton (2), and it was the implication raised by these findings that led to the exploratory study that is presently to be described.

B. PROCEDURE

The study was carried out in July of 1963 to determine the effect on an observer of his commitment to a political party or to a particular party candidate. The dependent variable was the bias in the observer's evaluations as debaters of his own and the opposing party's presidential candidate. Forty-nine persons took part in the study (26 males and 23 females). Ages ranged from 18 to 52 years, the median was 29. Sixty-three per cent of the sample had at least three years of college, and 27 per cent had some graduate training. The sample included professional people, clerical and research assistants, administrative people, police officers, and undergraduate college students.

An experiment was carried out on independent groups consisting of two to four persons. *Ss* were seated beside one another in a single row separated laterally by approximately two feet and faced a tape recorder. A response apparatus was fastened to the arm of each *S's* chair. The response apparatus was enclosed in a metal box (about the size of a shoe box) into which *S* could place his hand (thus shielding his responses from his neighbors) and grasp a lever that could be moved in a horizontal arc. *Ss* were instructed that moving the lever all the way to the left would register an extreme negative reaction; moving it all the way to the right, an extreme positive reaction; and the center position would register neutral. The lever was spring loaded to return to the center position. Though *Ss* were led to believe that their responses were scored on a continuous scale for intensity of feeling, responses were in fact recorded on a discrete scale with five categories of judgment, for each of which the duration of reaction time was independently recorded. Reaction

was measured continuously in time concurrent with continuation of the stimulus, and total time was checked by comparing the summed reaction time for each *S* against an independent total-time counter. This experimental apparatus was a modification of the Lazarsfeld-Stanton Program Analyzer. Where the Program Analyzer records a dichotomous response—like, dislike—the present apparatus permitted a continuous and variable intensity automated response: hence, its name, Caviar.²

After *Ss* had been seated and introduced to *E*, each *S* completed an 11-item paper-and-pencil test designed to measure the intensity of his political party commitment and his relative commitment to the two major presidential candidates for 1960. A consistent attribution of opinions and preferences in a direction favoring either one political party or one political candidate was taken as indicating, respectively, high commitment to party or high commitment to candidate. Inconsistency in the direction of opinions and preferences was taken as indicating either commitment lower in the individual's response hierarchy or no commitment at all, depending upon the extent of the inconsistency.

Following completion of the questionnaires, *E* told *Ss* that they were about to hear excerpts from the second Nixon-Kennedy debate of October 7, 1962. They were instructed to listen carefully both to the questions asked each of the candidates as well as to the candidates' answers. They were told to disregard their own personal feelings and to rate the "quality" of the candidates' answers, judging these on the basis of logic, clarity, and conciseness. *Ss* were instructed to regard the overall quality from a debater's point of view, considering whether or not the candidates' answers hit home, whether or not they struck a telling blow against the opposition, or whether or not the answers were longwinded, hollow, unclear, or perhaps more laden with emotion and demagoguery than with sound logic. Did the candidate's statements leave the debater in an indefensible position? Was he now vulnerable to attack by the opposition? And so on.

After a brief opportunity to practice with the response apparatus, *Ss* listened to an edited tape recording of the second Nixon-Kennedy debate. The tape was edited in a manner such that one candidate responded to four questions, followed by a rest period of approximately three minutes (permitting *E* to reset the recording apparatus), after which time the answers to four questions by the other candidate were presented. The tape was approximately 15 minutes

² The present version of the Program Analyzer was developed by the Bureau of Social Science Research, Inc., under contract SD-IE-7, Directorate for Armed Forces Information and Education, Department of Defense.

long, with about an equal amount of time allotted to each candidate. In order to balance out any complicating order effect, half of the *Ss* heard Nixon first; while the other half heard Kennedy first.

Ss' responses in each of five categories of judgment—"extremely good," "somewhat good," "neither good nor poor," "somewhat poor," "extremely poor"—were independently recorded on an apparatus located in a room adjacent to the experimental room. Continuous measurements were taken on each *S*, the duration of responses in each of the categories measured to an accuracy of .1 second. To simplify later statistical computation and to clarify the meaning of scores, *Ss'* responses were converted to a five-point Likert-type scale; that is, the response "extremely good" was considered as a numerical score of five, "somewhat good" as four, and so on. Two evaluation indices (*EI*) were calculated for each *S* from his independent ratings of the candidates, yielding an overall rating for each *S* of the quality of each of the candidates as debaters. The evaluation index was adapted from the median formula for grouped data. *EI* was defined as a measure of central value on the five-point scale such that above or below that measure the duration of reaction time was 50 per cent of the total reaction time. The general formula was as follows: $EI = 2.5 + [\Sigma t/2 - (t_1 + t_2)]/t_3$ where 2.5 is the lower limit for one interval in which the central measure usually fell. The symbols t_1 , t_2 , and t_3 represent categories "extremely poor," "somewhat poor," and

TABLE 1
ANALYSES OF VARIANCE FOR COMMITMENT

Source	<i>SS</i>	<i>df</i>	<i>F</i>
High commitment to party <i>vs.</i> low commitment to party	.046	1	.393
Nixon <i>vs.</i> Kennedy	.525	1	4.487*
Interaction			
Party commitment $\times$ Candidates	.179	1	1.530
Within cells	7.025	60	
Total	7.775	63	
High commitment to candidate <i>vs.</i> low commitment to candidate	.010	1	.084
Nixon <i>vs.</i> Kennedy	.114	1	.958
Interaction			
Candidate commitment $\times$ Candidates	.075	1	.630
Within cells	7.134	60	
Total	7.333	63	

* $p < .05$.

"neither good nor poor." Σt represents the total reaction time, summed for the five categories.

C. RESULTS

Out of the sample of 49 persons, a cluster of 23 scored high on party commitment, and a cluster of nine scored low. The scores were significantly different from one cluster to the other ($p < .05$), and a two-factor analysis of variance was computed from the *EI* scores for these people. Each factor consisted of two levels—one, high commitment and low commitment to party; the other, ratings of Nixon and ratings of Kennedy. These results are summarized in Table 1. In addition, a two-by-two analysis of variance was computed from the *EI* scores for significantly different clusters of *Ss* scoring high and low on commitment to candidate. These results are also summarized in Table 1. The relationship between party-commitment and candidate-commitment scores was determined by computing a Pearson product-moment correlation on these pretreatment questionnaire responses. The scores were not significantly related to each other, $r = .139$.

D. DISCUSSION

The two analyses of variance summarized in Table 1 indicate a single effect significant (beyond the .05 level) for *Ss* pooled on the basis of their intensity of party commitment, but nothing for *Ss* pooled on the basis of candidate commitment. This finding is particularly interesting because from a common-sense point of view one could argue that high commitment to party or high commitment to candidate or both could relate to a person's being biased in his judgment of candidate performance in a debate. Despite this common-sense contention, the data suggest that party commitment might operate to bias ratings of candidates; whereas candidate commitment does not.

This finding is all the more interesting when one realizes that despite some overlap between the two subsamples, the overlap is minimal; that is, some *Ss* who were selected on the basis of party commitment were also selected for their candidate-commitment scores. Moreover, not only are these subsamples peopled to an extent by different individuals; but the variables that are the basis for forming the subsamples bear an almost negligible relationship to one another ($r = .139$). The fact remains, however, that these different subsamples did rate the debaters differently; yet the hypothesized effect for interaction is not significant.

To get further insight into the problem, subgroups were formed on the basis of *Ss'* responses to an item that had asked, "If you had to be affiliated with one of the following, which would you choose? Democratic Party, Re-

publican Party, Other." Table 2 presents the relationship of major party affiliation and evaluation of debaters. Chi square, 16.45 (with two degrees of freedom), is significant at the .05 level. It is apparent that belonging to or

TABLE 2
RELATIONSHIP OF PARTY AFFILIATION TO EVALUATION OF QUALITY OF DEBATERS

Relative rating of candidate	Rater's Party Affiliation	
	Republican	Democrat
Nixon rated better than Kennedy	14	2
Nixon and Kennedy rated as equal	8	5
Kennedy rated better than Nixon	2	12

being "affiliated with" a political party bears some relationship to the affiliated individual's perception of his own party's presidential nominee. It is interesting to note that this across-the-board relationship is much different from the common-sense relationship originally posited. Evaluative bias was not confined simply to highly committed people, but was distributed throughout the sample. This suggests that some other independent variable may be responsible for the effect. However, equally as interesting is the finding in Table 2 that neither party seems to hold a monopoly on biased perception. The number of biased Democrats was about equal to the number of biased Republicans. When we take a closer look at these 14 Republicans and 12 Democrats, we find that they are about as heterogeneous en masse as was the overall sample from which they were drawn. There are professional people, clerical and research assistants, some administrative people, police officers, and the ubiquitous undergraduate college student. Fifteen of the 26 are male, age ranges from 18 to 52, with a median age of 29. Despite the small size of the sample and the exploratory nature of the study, there remains the implication that the 28 persons may not be so unlike some that may be found in a national survey on the effects of some future presidential debate.

To review, the study demonstrates a significant bias in the way politically affiliated people perceive opposing party presidential nominees. Despite having been instructed to the contrary, about an equal number of both Democrats and Republicans demonstrated this judgmental bias. Most important was the finding that the bias could not be related to a single common-sense independent variable—high commitment. Rather, the effect was noted in a high proportion of Ss almost across the board.

A study that purports to infer the effects of the presidential debates from the "objective" reactions of observers must certainly take steps to ameliorate the judgmental bias that has been demonstrated. The point is that the effects

produced by the debates will not be merely a function of the stimulus complex (the debates) but will in large measure be related to the characteristics of the judges. Consequently, in order to get any evaluative measure by judges about the debates, a great deal of descriptive information about the judges may be required. The fact that they were randomly selected and that their judgments were independent of one another may not be enough. Randomization, however, does give some assurance that the judgmental error is at least loaded comparably into both subsamples of Democrats and Republicans, assuming that the size of the subsamples is about equal. Perhaps a more satisfactory technique for making allowance for this uncontrolled variable is to determine the sampling-error adjustment that needs to be made to correct for judgmental bias. In the case where one may assume that ratings (or perhaps objective reports of the effects, or behavioral measures of the effects, etc.) of the debates are dispersed normally about a central measure, then the analysis of covariance should be considered, but what the particular covariance variables must be would require additional study.

One other point needs to be made, and that is that though high commitment to candidate yielded no significant effect, the situation might have been very different were the debates of the preconvention intraparty type—that is, Democrat versus Democrat, or Republican versus Republican—where candidate commitment might have been even more salient for the Ss. Moreover, in the present situation, Ss' involvement was probably considerably more reduced than would it have been in 1960, and this suggests that raters might have been influenced in some measure by the prestige of Kennedy's office. Support for this interpretation is found in Table 3, which presents the relationship

TABLE 3
RELATIONSHIP OF RELATIVE CANDIDATE COMMITMENT TO PARTY AFFILIATION

Affiliation preference	Relative Candidate Commitment		
	Nixon	None	Kennedy
Republican Party	12	5	7
Democratic Party	1	3	15

between ordinally scored candidate commitment and party-affiliation preference. With two degrees of freedom, $\chi^2 = 12.22$, and is significant at the .05 level. The datum suggests solid commitment from the Democrats, but a "splintering" of the Republicans approximately one-third of whom show a greater measure of commitment to Kennedy than to their party's nominee. This finding is reminiscent of the sleeper effect described by Hovland and Weiss (4) in their well-known study on communicator credibility. That is,

had the present study been carried out closer in time to the original debate the splintering might not have been as great. One might hypothesize that where an immediate posttest would have picked up a sharp division on candidate commitment between the two major parties, but solid commitment within each party, a delayed posttest would show a tendency in the losing party toward a splintering effect.

This "splinter" effect might be explained in the context of Festinger's theory of cognitive dissonance (3). In the case of the individual initially committed to a group or idea that has been granted irrevocable and public acceptance by a majority of other persons, such commitment would fail to subject the individual to any cognitive stress. Consequently, he would experience no motivational force to alter his initial commitment. However, in the case of the individual who is initially committed to a group or idea that has been irrevocably and publicly rejected by other persons, the individual's commitment is dissonant with the cognition of that rejection. Such an individual would experience strong cognitive pressures to reduce that dissonance, and the dissonance might be reduced by his repudiating his initial commitment and becoming more committed to the accepted group or idea.

E. SUMMARY

An experiment was carried out to determine the role of "commitment" to a political party or to a particular party nominee and the subsequent biasing effect on evaluation of the presidential debates. Forty-nine persons were subjected to measurements of the intensity of such commitments, thereafter exposed to excerpts from the 1960 presidential debates while concurrently rating the debaters using a modified Program Analyzer. Despite their having been instructed to evaluate "dispassionately" the debaters, bias in judgment was obtained. The bias was not, however, confined to persons with high commitment to party or high commitment to candidate, but was noted in a high proportion of Ss almost across the board. Suggestions for ameliorating such judgmental error were discussed.

In addition, a "splinter" effect was reported. That is, two and one-half years after election of the Democratic candidate, it was found that while the Democratic Ss remained fundamentally unchanged in their commitment to the Democratic nominee, approximately one-third of the Republican Ss manifested greater commitment to the President than to their own party's nominee. This "splintering" of commitment was explained in the context of cognitive-dissonance theory.

REFERENCES

1. BASS, B. M., & DUNTEMAN, G. Bias in the evaluation of one's group, its allies and opponents. *J. Confl. Resol.*, 1963, 7, 16-20.
2. BLAKE, R. R., & MOUTON, J. S. Competition, communication, and conformity. In I. A. Berg & B. M. Bass (Eds.), *Conformity and Deviation*. New York: Harper, 1961. Pp. 199-229.
3. FESTINGER, L. A Theory of Cognitive Dissonance. Stanford: Stanford Univ. Press, 1957.
4. HOVLAND, C. I., & WEISS, W. The influence of source credibility on communication effectiveness. *Publ. Opin. Quart.*, 1951, 15, 635-650.

Communication Research Division
Boston University
640 Commonwealth Avenue
Boston, Massachusetts 02215

SOCIAL-GROUP MEMBERSHIP, LEADERSHIP,
AND BIRTH ORDER*

Addiction Research Foundation, Toronto, Ontario, Canada

REGINALD G. SMART

A. INTRODUCTION

Several attempts have been made to relate affiliative behavior to birth order: for example, Schachter (8) has shown that the need for affiliation during anxiety is greater among first-born children than among later-born children. Becker and Carroll (1) have suggested that first-born children's greater need for affiliation leads them to join more social and recreational groups and that most groups contain a preponderance of first-born children in their membership. The possibility that first-born children have more social-group memberships than later-born children deserves study for several reasons. The first is that such study could extend our knowledge of the background correlates of affiliative behavior; the second derives from our knowledge that first-born children, compared to later-born children, demonstrate different personality traits, such as greater dependency in school and peer relationships (4); stronger responses to frustrating acts by others (3); and more tendencies to anger, intense emotionality, and violent reactions to defeat (5). This raises the possibility that small-group studies or experiments employing members of social groups as subjects involve biased samples in terms of birth order and many of the characteristics associated with birth order.

Becker and Carroll (1) have described some informal data (from a study of delinquent gangs) that they think imply that first-born children are over-represented among gang members. Unfortunately, Becker and Carroll's data, meant to be only suggestive, do not indicate the expected proportions of first-born, later-born and intermediate-born children under the null hypothesis of no birth-order difference. A complicating factor is suggested by Sampson's study (7), which showed that first-born females are *more* resistant to social influence than are later-born females; however, first-born males are significantly less resistant to social influence than are later-born males. A problem with Sampson's study is that the difference for females is significant only at the .10 level, and that level usually is *not* taken as evidence for the support

* Received in the Editorial Office, Provincetown, Massachusetts, on May 4, 1964. Copyright, 1965, by The Journal Press.

of an hypothesis; so, among females, later-born children should have numerically but not significantly more group memberships than first-born children. Among males, the opposite should be true, with first-born children having significantly more group memberships than later-born children. One purpose of this paper is to present data relevant to the foregoing hypotheses.

The possibility that birth order and democratic leadership in groups are related has apparently never been examined, although several investigations suggest such a relationship. If, as some studies indicate (6), group leaders reflect the mean attitude of the group with regard to conformity, then first-born children should be more often elected as leaders than later-born children. It is easy to see how the first-born (hence, oldest) child would show leadership among his siblings, although it might be more autocratic than democratic. In addition, Dean (2) and Koch (5) have found that first-born children do show more leadership when young. This study predicts that first-born group members are more likely to hold elected offices in their groups as a reflection of their greater leadership characteristics.

B. METHOD

The data for this study were obtained from questionnaires distributed to 370 students (197 males and 173 females) enrolled in four undergraduate courses at the University of Toronto. The students were asked to list the number of older and younger siblings: living as well as deceased. Siblings were taken to contribute to the student's ordinal position only if they had lived for the first five years of the respondent student's life. The students were also asked to list the social and recreational clubs to which they belonged, together with any elected offices they held. Social or recreational activities per se were not counted unless they occurred in connection with some formal club membership. In addition, appointed positions were not counted as evidence of leadership.

C. RESULTS

All subjects were categorized as only children, first-born children, intermediate-born children, or last-born children. Further, each student was categorized as having no club memberships or one or more memberships. The data from first-born children and only children were combined to form a first-born; those for intermediate-born children and last-born children were combined into a later-born group.

The numbers of first-born males and later-born males who had one or more

TABLE 1
THE NUMBER OF FIRST-BORN MALES AND LATER-BORN MALES
AND GROUP MEMBERSHIP*

Birth order	Group memberships		Total
	One or more	None	
First born	57	34	91
Later born	51	55	106
Total	108	89	197

* Chi square = 4.239, $df = 1$, $p < .05$.

group memberships and those who had no memberships is shown in Table 1. The chi square, 4.239, is significant at the .05 level and shows that more first-born males are group members than are later-born males. Comparable data for females are shown in Table 2 and, in keeping with expectation, later-born females are more often members of social groups than are early-born females, but the difference is not significant (chi square = .205, $p > .80$).

TABLE 2
THE NUMBER OF FIRST-BORN FEMALES AND LATER-BORN FEMALES
AND GROUP MEMBERSHIP*

Birth order	Group memberships		Total
	One or more	None	
First born	38	54	92
Later born	37	44	81
Total	75	98	173

* Chi square = .205, $df = 1$, $p > .80$.

Each student who had one or more group memberships was characterized as holding or not holding an elected office. Since no differential prediction based on sex was made, the data for males and females were combined. The numbers of group members who held office and who did not hold office are shown in Table 3. Since the chi square is nonsignificant (chi square = .696, $p > .80$) there is no indication that leadership is associated with birth order.

TABLE 3
THE NUMBER OF GROUP MEMBERS WHO HELD AND WHO DID NOT
HOLD ELECTED OFFICE*

Birth order	Held elected office		Total
	Held elected office	Did not hold elected office	
First born	29	66	95
Later born	22	66	88
Total	51	132	183

* Chi square = .696, $df = 1$, $p > .80$.

D. DISCUSSION

The results support both predictions for affiliative behavior. First-born males more often than later-born males are social-group members, but there is no birth-order difference for females. This finding adds further support to the proposition that first-born males have greater affiliative needs than do later-born males and supports Sampson's findings (7) and Becker and Carroll's expectations (1). This result suggests that social groups composed wholly or mainly of males contain a preponderance of relatively dependent persons who are sensitive to social influence. This assertion, which of course bears further investigation, could seriously limit the generality of many socio-psychological investigations using members of social groups.

The failure to obtain a significant birth-order difference for females is in line with the expectation based on Sampson's findings (7). Affiliative behavior seems to be associated with birth order only in males; however, later-born females might demonstrate affiliative behavior in ways different from those of first-born males. Fewer females than males, regardless of birth order, were members of social groups; and it may be that female affiliative needs are more often met by informal social groups. Perhaps greater birth-order differences would be found for females if questions were asked about the number of evenings spent alone, the number of times they go to classes alone, or about other less-formal social events.

The independence of birth order and leadership in groups is contrary to the prediction based on studies of young children. Perhaps the first-born child's early leadership behavior is lost by the late teens or perhaps their leadership is more autocratic than democratic—as implied by holding elected offices. Further studies of the development of leadership tendencies and of different types of leadership are needed before these questions can be satisfactorily answered.

E. SUMMARY

The aim of this study was to relate social-group membership and leadership to birth order. Based on studies of conformity and affiliative behavior it was predicted that first-born males have more social-group memberships than do later-born males, but no significant birth-order difference was expected of females. First-born children were expected to be more often chosen for elected offices (indicative of democratic leadership) than are later-born children. These hypotheses were examined by collecting relevant data from undergraduate students. It was found that first-born males do have more social-group memberships, whereas no difference obtains for females. The predicted dependence of leadership on birth order was not found.

REFERENCES

1. BECKER, S. W., & CARROLL, J. Ordinal position and conformity. *J. Abn. & Soc. Psychol.*, 1962, **2**, 129-131.
2. DEAN, D. A. The relation of ordinal positions to personality in young children. M.A. thesis, State University of Iowa, Iowa City, 1947.
3. DITTES, J. Birth order and vulnerability to differences in acceptance. *Amer. Psychologist*, 1961, **16**, 358 (Abstract).
4. GEWIRTZ, J. Dependent and aggressive interaction in young children. Ph.D. dissertation, State University of Iowa, Iowa City, 1948.
5. KOCH, H. Some personality correlates of sex, sibling position, and sex of sibling among five and six year old children. *Genet. Psychol. Monog.*, 1955, **52**, 3-50.
6. MARSH, C. P., & COLEMAN, A. L. Farmer's practice adoption rates in relation to adoption rates of leaders. *Rural Sociol.*, 1954, **19**, 180-183.
7. SAMPSON, E. E. Birth order, need achievement, and conformity. *J. Abn. & Soc. Psychol.*, 1963, **64**, 155-159.
8. SCHACHTER, S. *The Psychology of Affiliation*. Stanford: Stanford Univ. Press, 1959.

Addiction Research Foundation
24 Harbord Street
Toronto 5, Ontario, Canada

EFFECT OF DISCREPANT COMMUNICATIONS ON SELF-PERCEPTION*¹

Department of Psychology, University of Texas

DEMETRIOS PAPAGEORGIS AND BRIAN M. McCANN

A. INTRODUCTION

It is becoming increasingly apparent that the variables studied in opinion and attitude change may be especially relevant to the study of personal-influence processes that characterize in part counseling and psychotherapy. In the past, authors of various psychotherapeutic systems and procedures that range from advice to prolonged intensive psychotherapy have tended to rely on "clinical" intuition and observation in order to understand their successful and unsuccessful attempts at communication. Some unfortunate consequences of this, at best, partial approach have been discussed by Eysenck (5). More recently, conceptualizations of psychotherapy as a learning process (2) and direct attempts to relate the learning and acceptance of self-relevant information (3) have appeared in the literature. In addition, several practitioners [e.g., Ellis (4), Ruesch (12a), Wolpe (14)] are now relying in various degrees on methods of treatment directly related to experimentally studied variables of learning and of attitude change. Furthermore, Frank (8) has written a well-documented account of the close connection between persuasion and personality change.

Of special relevance to the personality-change area is the effect of persuasive communications aimed at changing a person's existing perceptions of his personality and self. Bergin (3) found an interaction between communicator credibility and discrepancy between communication position and the subject's prior self-appraisal on the masculinity-femininity dimension. In this study, an extremely negative message produced most change if the communicator was of high credibility and least change if his credibility was low. Although Bergin's results provide an excellent demonstration of a prediction based on dissonance theory (6), it seems desirable that other personality traits and other communication variables should also be investigated. The experiment

* Received in the Editorial Office, Provincetown, Massachusetts, on May 7, 1964. Copyright, 1965, by The Journal Press.

¹ The research reported in this paper was supported in part by a grant to the senior author from The Hogg Foundation for Mental Health, The University of Texas.

to be reported introduces another personality dimension and attempts' also to investigate the efficacy of positive communication as well as that of negative communication.

Although the physical presence of a communicator is a standard procedure in counseling and psychotherapy, the experimental assessment of the precise communicator effect presents difficulties. The use of two different communicators [see Bergin (3)] appears to be an adequate solution, but it should be noted that a perceived difference in credibility is only one of a great number of ways that two people, physically present, may be perceived as different from each other. On the other hand, the presence of the same communicator acting different parts creates perhaps more serious difficulties in an experimental design. A possible step toward an unambiguous demonstration of communicator effects is to remove the physical stimulus of the communicator and to insure that, at least as distal stimulus, the source of the communication remains fairly constant for all subjects.

The study reported here was designed to investigate the effects of a discrepant communication on a person's self-perception, holding the source of the message as constant as possible. A "personality" test, suitable for group administration and self-scoring, was so constructed that the vast majority of subjects would obtain a very "high" score on one scale. For some subjects, the high-scoring scale was labeled "hostility"; while for other subjects, the high-scoring scale was labeled "masculinity" (for male subjects) or "femininity" (for female subjects). Thus, by means of a test described as standardized and valid, an attempt was made to hold the credibility of the source constant for all subjects; also to assess the effect of self-relevant communications of both negative and positive type.

Specifically, the authors predicted (a) that a discrepant communication of the type just described results in self-evaluations in line with the direction of the communication, and (b) that a negative or less socially desirable communication finds greater acceptance than does a positive communication. Prediction (b) was derived from dissonance theory (6), since a negative communication is expected to create greater dissonance (hence, greater pressure to reduce the dissonance). Alternatively, it may be argued that self-evaluations of positive characteristics are more resistant to change than are similar evaluations of one's negative characteristics. In all likelihood, the latter are seldom included within the self-concept; consequently, if attitudes about negative characteristics are relatively hazy and uncrystallized, they may be more vulnerable to persuasive communication.

It should be pointed out that even though the relevance of the present ex-

periment to some aspects of counseling and psychotherapy seems obvious, it is also true that by no means does it come close to an experimental analogue of either. Rather, the study reported represents a highly restricted investigation of a limited number of variables that may operate in communication situations when changes in self-perception are advocated.

B. METHOD

An attempt was made to change the subjects' self-perceptions of a positive personality characteristic (masculinity or femininity) or a negative personality characteristic (hostility) by means of an "extreme" communication concerning one of them. The experimental groups were given a self-scoring group "personality" test so designed that each subject would score close to average on three test scales and extremely high on a fourth scale. This fourth scale, the Deviant Scale (D), was labeled "masculinity," "femininity," or "hostility" depending on the experimental condition. The post-test self-ratings of the experimental groups on the manipulated characteristic were compared to those of the control group who had also been given the test but who had not scored it; hence, received no communication. Thus, the design conformed to the so-called after-only design [Hovland, Lumsdaine, and Sheffield (10)] averting the danger of precommunication sensitization and enhancing the likelihood of acceptance of the stated purpose of the experiment at face value.

1. Subjects

Eighty-eight male and 78 female undergraduates enrolled in an introductory-psychology course at The University of Texas served in the experiment. While all students in the course are required to serve as experimental subjects, participation in any given experiment is voluntary. Groups of subjects were randomly assigned to one of the following experimental or control conditions: (a) masculinity deviant ($N = 30$) or femininity deviant ($N = 22$), (b) hostility deviant (30 males and 26 females), (c) control group (28 males and 29 females); i.e., subjects who took the test but did not score it.

In addition, 17 male subjects and 11 female subjects drawn from the same student population supplied semantic-differential ratings [see Osgood, Suci, and Tannenbaum (12)] of the characteristics used in the communications in order to establish how these were understood and evaluated among the students.

2. Materials

A 107-item mimeographed "personality" test composed of items taken from the Minnesota Multiphasic Personality Inventory was given to the two experi-

mental groups and to the control subjects. Answers to the test were recorded by each subject on an answer sheet. Later, the experimental subjects received four specially constructed scoring stencils and a profile sheet on which they were to plot their scores from the four scales. The scales consisted of the Lie Scale (L), the Correction Scale (K), and the Psychopathic Deviate Scale (Pd) taken intact from the MMPI and scored in the usual way, and the contrived Deviant Scale (D). The last scale consisted of 25 MMPI items that had been answered in one direction only by more than 95 per cent of the males and females in the MMPI college-standardization sample. The scoring stencil for the contrived D Scale was keyed, so that on 21 of the 25 items the most-common response contributed to the total raw score for the scale. The authors expected the vast majority of the subjects to score high on the D Scale and close to the mean on the remaining scales.

The profile sheet was a simplified MMPI profile reproduced by mimeograph. At the bottom of the profile were the names of the four fictitious scales, appearing in a constant order for all experimental groups and called "Masculinity" or "Femininity," "Creativity," "Social Adequacy," and "Hostility." All subjects were instructed to write the score obtained from the MMPI Lie Scale under "Creativity" and that obtained from the Correction Scale under "Social Adequacy." For the masculinity-deviant and femininity-deviant groups, the D Scale scores were labeled "Masculinity" or "Femininity"; the Psychopathic Deviate scale scores, "Hostility." The reverse of this labeling was used with the hostility-deviant groups. As with the MMPI profile sheet, raw scores on each of the four scales were located on the center of the sheet; and the corresponding centile scores were determined from a conversion scale at each side. A solid dark line drawn at the 50th centile point indicated the mean score for the "test," while dotted lines along the 35th and 65th centiles were labeled as including the "average" range of scores for each personality characteristic.

The experimental measure consisted of self-ratings of five personality characteristics described as self-satisfaction, masculinity or femininity, number of problems, hostility, and adjustment. Nine statements were presented for each characteristic, and the subject rated himself in comparison to the "average person" ("average male" and "average female" were substituted in the case of masculinity and femininity). The statements ranged from "I am extremely more ——— than the average person" (Score 9) to "I am extremely less ——— than the average person" (Score 1). Each subject was asked to indicate what statement he considered most acceptable and most objectionable to him as well as his latitudes of acceptance and rejection [see Hovland, Harvey, and Sherif (9)].

3. Procedure

Subjects, randomly assigned to hostility-deviant, masculinity-deviant or femininity-deviant, and control-group conditions, were tested in mixed-sex groups of 10 to 20. The subjects were spread out in a large classroom in an attempt to preclude their seeing any test results but their own because it was anticipated that the test scores would be similar and surprising.

Subjects were told that they were being tested as part of a nationwide project designed to determine some personality characteristics of American college students. Both the contrived test, which was named the "California Polyphasic Personality Inventory" and the experimental measure that followed were presented as part of this testing. The former was described as a valid, standardized, and reliable instrument for which national norms were available. The students were also told that each test item had predictive validity, even though some items lacked face validity (a brief explanation of the technical terms was also given). When all the experimental subjects in a group had finished marking their answer sheets, the test booklets were collected; and the scoring stencils and profile sheets were distributed with instructions for scoring, plotting, and interpreting centile scores. This part of the procedure was omitted in the control groups.

The second "test" (the experimental measure) was given after the subjects had been instructed to put aside the other materials. Detailed instructions were given.

Each group session was supervised by the same two experimenters (one, a male psychology faculty member; the other, a male graduate-student assistant). Instructions were always given in a standard fashion by the first-mentioned experimenter. The total time required for experimental groups averaged 35 minutes. The subjects responded anonymously.

At the end of the experiment, the subjects were informed of the true nature and purpose of the study, and were especially cautioned that the results obtained on the contrived test were in no way related to their actual standing on the four characteristics.

C. RESULTS

1. *Semantic-Differential Ratings of the Characteristics*

It will be recalled that semantic-differential ratings of the characteristics used in the experiment were obtained from 28 male and female subjects (drawn from the same student population), who did not participate in the experiment proper. The main purpose of these ratings was to verify the assumption that hostility represented a negative trait and that masculinity

and femininity were viewed positively. The subjects were asked to rate several concepts including "hostility," "a man who is very hostile," "a girl who is very hostile," "masculinity," "a man who is very masculine," "femininity," and "a girl who is very feminine." The concepts were rated on 15 bipolar scales, five of which represented the dimension of understandability [see Nunnally (11)], with another five based on the probabilistic or "belief" dimension, and the remaining five relevant to the evaluative or "attitude" dimension [see Fishbein and Raven (7)].

The results of immediate interest were obtained on the evaluative (e.g., good-bad) dimension of the concepts. Hostility and its related concepts received uniformly low evaluations from both male and female subjects, the mean scores ranging from 2.85 to 3.67 on the seven-point semantic-differential scale. The concepts related to masculinity and femininity uniformly received positive evaluations, the mean ratings ranging from 5.13 to 5.87. The mean differences between evaluations of hostility and masculinity or femininity concepts are significant at the .01 level; thus supporting the assumption that the experimental subjects did indeed receive positive communication (masculinity or femininity) or negative communication (hostility).

Differences similar to those found with the evaluative ratings were also observed on the understandability dimension. Here, too, the hostility concepts were rated as significantly less understandable ($p < .01$), less familiar, and more complex than either masculinity or femininity. Finally, the belief dimension (e.g., possible-impossible) yielded uniformly high scores on all concepts, with a slight tendency for the positive concepts to receive higher ratings. On all three meaning dimensions sex differences were observed, but they were not consistent. If anything, female subjects tended to rate lower than males.

2. *Success of the Experimental Manipulation*

The experimental subjects plotted their scores from the four scales on the contrived self-scoring test on a profile sheet. It was intended that each subject would receive a high score on the D Scale and approximately average scores on the remaining three scales. Since the experimental manipulation was to some extent under the subjects' control, it is necessary to show that the experimental subjects did, as was intended, receive discrepant high scores on the test profile. Only one female subject, who misused the scoring stencils, received a score below the 65th centile on the D Scale; so she was excluded from the results summarized below. The mean centile scores for the various experimental groups ranged from the 84th to the 86th centile. In contrast, the corresponding mean centile scores for the same characteristics in the nondeviant communication conditions (i.e., from the MMPI Psychopathic

Deviate Scale) ranged from the 35th to the 39th centiles, below the mean but still within the average range indicated on the profile. Average scores were also obtained on the remaining two noncritical scales.

3. *Effect of Communications on Self-Ratings*

Following the scoring of the test and the plotting of the profile, or taking the test only for the control subjects, the students rated themselves on personality characteristics, including those experimentally manipulated. These ratings allowed comparisons of experimental and control subjects on what statement they chose as most acceptable and most objectionable, as well as on measures of their latitudes of acceptance and rejection. The latitude of acceptance was defined as the most-acceptable statement plus all other statements marked as acceptable; the latitude of rejection, as the most-objectionable statement and all other statements considered objectionable. Two measures were computed for each latitude: its size or the number of statements it included and its mean.

As expected, the most-acceptable-statement rating proved to be the most useful of the foregoing indices [see Bergin (3)]. The results for this measure with experimental and control groups are shown in Table 1. None of the mean

TABLE 1
MOST-ACCEPTABLE SELF-RATINGS FOR THE EXPERIMENTAL AND CONTROL GROUPS
ON THE TWO PERSONALITY CHARACTERISTICS

Parameter	Masculinity deviant (<i>N</i> = 30)	Femininity deviant (<i>N</i> = 22)	Male hostility deviant (<i>N</i> = 30)	Female hostility deviant (<i>N</i> = 25)	Male control (<i>N</i> = 28)	Female control (<i>N</i> = 29)
Masculinity or femininity						
Mean	6.47	5.32	5.43	5.04	5.75	5.34
<i>SD</i>	1.41	1.52	1.07	1.06	1.17	0.85
Hostility						
Mean	4.17	3.95	5.43	4.96	4.04	4.14
<i>SD</i>	1.78	1.56	1.68	1.69	1.45	1.38

ratings of the characteristics that received average profile scores differs significantly from the corresponding mean ratings obtained from the control subjects. According to Table 1, ratings of hostility by the masculinity-deviant group are not significantly different from those made by the masculinity-control group, the same being true of the femininity-deviant and femininity-control hostility ratings. Likewise, masculine hostility-deviant subjects do not differ from masculine-control subjects in masculinity ratings, nor do feminine hostility-deviant subjects differ from feminine-control subjects in femininity

ratings. Corresponding product-moment correlation coefficients between communicated centile score and most-acceptable self-rating are as follows: 0.50 ($p < .01$) for males and 0.47 ($p < .05$) for females on hostility, -0.14 (n.s.) for the masculine hostility-deviant group on masculinity, and $+0.20$ (n.s.) for the feminine hostility-deviant group on femininity. The discrepancy between the significant hostility correlations and the nonsignificant masculinity and femininity correlations is probably due to differences in the traits themselves. It is also possible that the hostility communications were closer to the subjects' initial position, while the masculinity and femininity communications advocated a more discrepant, or lower position, as may be seen from the control data.

Table 1 shows that not all communications were successful in producing higher self-ratings. Hostility-deviant communications for male subjects resulted in higher mean most-acceptable ratings than those expressed by the controls ($p < .01$). The corresponding communications for female subjects also resulted in higher ratings ($p < .06$), but the magnitude of the difference was smaller. Masculinity-deviant communications resulted in ratings that were slightly higher than those of the control-male subjects ($p < .05$), whereas femininity-deviant communications produced no significant changes in female subjects. The corresponding correlation coefficients between communication centile scores and most-acceptable ratings are as follows: male hostility-deviant subjects 0.12 (n.s.), female hostility-deviant -0.27 ($p < .01$), male masculinity-deviant subjects -0.11 (n.s.), and female femininity-deviant subjects, -0.02 (n.s.).

The other measures obtained from the experimental questionnaire did not yield results that were as meaningful as those just reported. In terms of most-objectionable ratings, only the male hostility-deviant subjects showed a significant tendency ($p < .01$, by Fisher's exact probability test) to reject more often the extreme low rating on the trait than did the controls. No significant differences were obtained in the sizes of the latitudes of acceptance and rejection, although the findings consistently suggested that the communications were effective. The mean latitude of acceptance and rejection scores contained several significant effects which followed the results obtained with the mean most-acceptable ratings. It should be noted that this last measure is not independent of the most-acceptable and most-objectionable ratings.

D. DISCUSSION

The first hypothesis, which predicted self-evaluations in line with the discrepant communication, receives unambiguous support only in the case of

male subjects. Female students also rated themselves higher than the controls after learning their deviant scores on hostility, but the level of significance ($p < .06$) is not high. A conservative conclusion is that the first hypothesis is supported for male subjects with both types of communication, while female subjects show a tendency to be affected only by the negative, hostility communication.

The second hypothesis, which predicted a greater degree of acceptance of the negative communication is supported in the case of females. No clear support is available for male subjects, although inspection of the means possibly indicates greater acceptance of the negative communication.

The authors assumed that a communication that indicates a high standing on hostility is negatively evaluated by the subjects and that a communication that indicates a high standing on masculinity or femininity is positively evaluated. This assumption received support from the evaluative ratings of the concepts obtained from subjects who did not serve in the main experiment; however, the positive-negative evaluative dimension was not the only dimension on which differentiation of the personality concepts was possible: a similar relationship was found on the understandability dimension. Thus, although the distinction between positively and negatively evaluated communications may account for the findings, it is also possible that the degree to which a concept is understandable to the subject or has specific and precise meaning for him determines his persuasibility. The less-understandable or less-specific concept could be more susceptible to influence [Asch (1), Festinger (6), Sherif (13)]. It remains for future research to elucidate the role of this component of meaning by comparing, under the same conditions, explicitly defined personality traits to trait names without explicit definition or supplied with only a vague definition. Attitude change may be more likely to occur when the object of the attitude is capable of plausible redefinition by the subject.

The difference in susceptibility between males and females observed in this experiment remains a matter of speculation. The lesser susceptibility of females may be a cultural sex characteristic or may indicate a greater likelihood for female subjects to discount test information. The authors favor the hypothesis that femininity occupies a central position in the self-concept of females in our culture and may represent an attitude that is not only a matter of concern to the individual but is also stable enough to resist change. In an early part of this paper the authors advanced the notion that positive characteristics in general form a more basic part of the self-concept than do negative characteristics. The relatively small change observed in the masculinity dimension is also in line with this hypothesis; so it will be of some interest to see if the

significant masculinity effect obtained in this study holds in future replications.

Another interesting question for future research concerns the effectiveness of the group-administration conditions used in the present study. It appears that group conditions, with the possibility of overt subject reaction, makes the discounting of the validity of the communication easier for the subjects. At least this seems to be the case when highly negative messages are used. In a subsequent experiment, 30 male subjects were given very high scores on femininity and 23 female subjects obtained very high scores on masculinity. The overt reaction of these subjects to the communication indicated that group administration is not suited to highly discrepant negative communications of this type. On the other hand, with less-extreme messages, group conditions may facilitate a perception that the subject is undergoing a *bona fide* personality assessment. The self-scoring procedure may contribute further to this perception.

E. SUMMARY

Male and female college students took and scored a contrived "personality" test that resulted in extreme scores on a positive personality characteristic (masculinity or femininity) or a negative personality characteristic (hostility). Subsequent self-ratings showed acceptance of both communications by the male students and a tendency by females to accept only the negative message.

REFERENCES

1. ASCH, S. E. *Social Psychology*. New York: Prentice-Hall, 1952.
2. BANDURA, A. Psychotherapy as a learning process. *Psychol. Bull.*, 1961, 58, 143-159.
3. BERGIN, A. E. The effect of dissonant persuasive communications upon changes in a self-referring attitude. *J. Personal.*, 1962, 30, 428-438.
4. ELLIS, A. Reason and Emotion in Psychotherapy. New York: Lyle Stuart, 1962.
5. EYSENCK, H. J. The effects of psychotherapy. In H. J. Eysenck (Ed.), *Handbook of Abnormal Psychology*. New York: Basic Bks., 1961. Pp. 697-725.
6. FESTINGER, L. A Theory of Cognitive Dissonance. Evanston, Ill.: Row, Peterson, 1957.
7. FISHBEIN, M., & RAVEN, B. H. The AB scales—an operational definition of belief and attitude. *Hum. Relat.*, 1962, 15, 35-44.
8. FRANK, J. *Persuasion and Healing: A Comparative Study of Psychotherapy*. Baltimore: Johns Hopkins Press, 1961.
9. HOVLAND, C. I., HARVEY, O. J., & SHERIF, M. Assimilation and contrast effects in reactions to communication and attitude change. *J. Abn. & Soc. Psychol.*, 1957, 55, 244-252.
10. HOVLAND, C. I., LUMSDAINE, A. A., & SHEFFIELD, F. D. *Experiments on Mass Communications*. Princeton: Princeton Univ. Press, 1949.
11. NUNNALLY, J. C., JR. *Popular Conceptions of Mental Health*. New York: Holt, Rinehart, & Winston, 1961.
12. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana, Ill.: Univ. Illinois Press, 1957.

- 12a. RUESCH, J. *Therapeutic Communication*. New York: Norton, 1961.
13. SHERIF, M. Group influences upon the formation of norms and attitudes. In E. E. Maccoby, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in Social Psychology* (3rd ed.) New York: Holt, 1958. Pp. 219-232.
14. WOLPE, J. *Psychotherapy by Reciprocal Inhibition*. Stanford: Stanford Univ. Press, 1958.

Department of Psychology
University of Texas
Austin, Texas 78712

DIAGNOSTIC USE OF THE INTERACTION PROFILE*

Harvard University and Haverford College

ROBERT F. BALES AND A. PAUL HARE

A. INTRODUCTION

Since the publication of Interaction Process Analysis [Bales (1)], the method described there for the observation of social behavior has been used by a number of different investigators. Experience has accumulated to the point at which some useful hints may be given to the investigator who has taken observations of only one or a few groups and wishes to know what may be inferred about their peculiar characteristics. The present article summarizes the results obtained by different investigators in such a way as to provide a frame of reference for this sort of diagnostic use.

B. A REFERENCE POPULATION OF PROFILES

The population of profiles obtained by investigators who have used the method in a standard way for a given sort of group, task, or unusual condition provides a frame of reference within which any one profile gains added meaning. One may understand more about the particular situation from which his profile was obtained by discovering what other kinds of situations have given similar or different profiles.

In its raw form, an interaction profile is a frequency count of the number of scores that a given observer has entered in each category of the total set used for observation. The present system uses 12 categories, as shown in Table 1. Because scoring proceeds continuously throughout the course of the interaction observed, the raw number of scores obtained depends heavily upon the length of time. If chronological time has been measured, rates per unit of time for individual categories or for the total may be computed. Total interaction rate on a base of chronological time is a meaningful measure. Unfortunately, actual time rates are seldom reported; consequently, for comparisons across studies one must fall back on rates that are computed as percentages in each category of the total number of scores for all categories. We have reduced all profiles used in the present study to percentage rates.

Similarities and differences in profiles are attributable in part to peculiarities

* Received in the Editorial Office, Provincetown, Massachusetts, on May 7, 1964. Copyright, 1965, by The Journal Press.

TABLE 1
A REFERENCE POPULATION OF INTERACTION PROFILES

Interaction-process analysis category	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21																					
Standard-case discussion	Harvard 1955-1958*	Social conversation	Slater and Bales (20)	Discussion of self-ratings	Couch (7)	Planning role playing	Borgatta and Bales (5)	Projective stories	Mills (14)	Playing board game	Olmsted (15)	Subjects under lysergic acid	Slater, Morimoto, and Hyde (21)	Subjects under lysergic acid	Lennard <i>et al.</i> (13)	Drinking brandy and beer	Takala <i>et al.</i> (25)	Therapy groups	Psathas (16)	Prison therapy groups	Winter (26)	Paranoid and depressed patients	Roberts and Strodtbeck (18)	Therapy groups	Evans (8)	Psychiatric-stress interview	Hare <i>et al.</i> (11)	Jury deliberation	Strodtbeck and Mann (24)	Husbands and wives	Strodtbeck (23)	Husbands and wives	Quade (17)	Labor-mediation sessions	Landesberger (12)	Boy leaders	Hare (10)	Children's doll play	Gruber (9)	Group-process discussion	Borgatta <i>et al.</i> (6)	
1. Shows solidarity	3.5	8.8	5.4	2.3	.8	2.0	9.5	3.0	.3	3.4	2.1	1.8	2.1	.1	1.2	2.0	.6	2.4	4.3	2.5	4.3																					
2. Shows tension release	7.0	17.7	12.7	6.0	7.3	6.1	30.4	13.8	9.7	7.4	3.4	3.5	5.6	.5	3.0	5.0	5.3	1.2	6.2	7.5	12.2																					
3. Shows agreement	16.9	9.0	9.5	9.3	18.4	13.7	10.5	7.9	6.7	16.3	5.6	8.0	5.8	.9	17.9	17.0	9.7	10.5	4.7	5.0	21.5																					
4. Gives suggestion	8.0	0	2.2	5.8	25.2	20.2	7.1	3.5	3.7	.7	2.0	1.8	1.1	.5	4.0	6.0	14.5	3.0	16.3	9.0	3.1																					
5. Gives opinion	33.4	17.4	26.0	27.3	17.5	25.7	16.7	32.5	32.7	26.2	29.8	19.5	19.8	16.7	26.1	22.0	19.2	24.5	7.7	3.5	22.8																					
6. Gives information	15.5	28.7	20.2	33.8	14.5	14.2	13.6	21.0	23.9	29.8	33.2	35.8	42.2	56.9	38.3	29.0	29.0	32.7	34.2	37.0	19.7																					
7. Asks for information	2.3	11.0	7.1	6.8	3.6	3.0	3.2	4.7	1.8	5.1	3.0	8.4	2.9	12.9	4.6	6.0	8.2	7.3	9.7	8.0	4.0																					
8. Asks for opinion	2.2	1.8	4.0	1.9	2.2	3.8	1.2	7.9	1.0	3.7	9.3	3.6	7.3	2.8	1.9	4.0	4.5	2.2	.5	1.0	1.9																					
9. Asks for suggestion	.7	0	.4	.7	2.2	1.8	.6	1.0	.3	0	.3	.1	1.1	0	.2	1.0	1.5	.3	.1	0	.4																					
10. Shows disagreement	7.9	2.1	5.4	.9	6.4	7.7	5.3	2.4	14.3	3.7	7.0	2.6	4.7	.6	2.0	5.0	3.1	7.5	4.4	3.0	3.4																					
11. Shows tension	2.0	3.0	6.7	5.1	1.5	1.0	2.0	1.0	.2	2.7	1.8	6.0	4.6	6.1	.6	1.0	2.2	1.7	2.5	15.5	4.8																					
12. Shows antagonism	.5	0	.5	.1	.4	.2	.4	1.3	5.4	.8	2.4	7.9	2.7	1.8	.2	2.0	2.0	4.7	9.2	6.5	1.6																					

* Bales and Slater (2), (3), Berger (4), Salisbury (19).

of observers. In deciding what should constitute a single profile in the reference population it has seemed best to take only one summary profile from each study, so that the variance over profiles should most fairly represent the variance attributable to differences of investigators as well as to the situations they observed.

An interaction profile may be tabulated for a given subject vis-à-vis one or more others or for any summation of subjects. According to diagnostic needs, one might want a reference population when the single profile is the profile of a single individual or he might prefer the summary profile of sets of individuals with similar characteristics, such as sex or age, or he might prefer the single profile to be the summary of all individuals in a given interacting group, regardless of differences in characteristics. An example would be a summary profile of therapist and patients together in a therapy group. Finally, a number of groups observed under similar conditions might be added together to produce the single profile. We have taken this final alternative.

A single profile in our reference population usually represents a whole study by a single investigator, but there is one exception. In the laboratory studies at Harvard, a standard diagnostic task has been used for a number of studies by different investigators. For present purposes, all of the Harvard studies using this task have been summarized in a single profile to avoid overrepresentation of this task in the reference population; thus the profile we have taken as the single case for the reference population is either a summary profile for a given study or a series of studies under the same conditions. This procedure averages out the effects of all sources of variance internal to a given study. The gain, however, is that usually we are able to relate differences in profiles to major features of the situation, task, type or condition of subjects that have been described by the investigator.

Table 1 presents 21 summary profiles available at the time of writing (1963) to make up a reference population. Each profile is shown as a column of percentage rates in each category on the base of the total number of acts observed in the given study. A new profile may be placed in relation to the reference population by comparing it with each of the 21 until one is found which it most closely matches, but this is a laborious procedure. It would be easier to locate the peculiarities of a new profile if one had some idea of how far above or below the mean a given category generally has to be before we are able to connect it with some definite peculiarity of the situation, task, or type or condition of subjects.

C. MEANS AND VARIANCES IN THE REFERENCE POPULATION

The method for determining diagnostic cutting points usually involves computation of the mean and variance of each of the variables in the profile, but the present case presents special complications. The value of a given variable on a given profile in our population is a percentage rate. Unfortunately, percentage rates are not normally distributed within a sample drawn from a population of percentage rates. For rates of the magnitude obtained with the present method, the lower the mean rate the more skewed is the distribution because variation below the mean is restricted by the lower boundary of a zero rate; but variation is not restricted to the same degree above the mean. Fortunately, there exists a transformation that tends to normalize distributions of percentage rates: namely, the arc-sine transformation [Snedecor (22)]. For the computation of the variance in each category, we have applied this transformation to the percentage rates.

Because the conversion of a profile of percentage rates into arc-sine equivalents involves appreciable additional labor, it is convenient for diagnostic use (once the variances have been computed and the points of practical interest above and below the mean have been located) to have a table that states the mean and cutting points in percentage form. Table 2 gives the mean for each category and shows the points of the distribution that fall one standard deviation and 1.96 standard deviations above and below the mean.

TABLE 2
MEAN PERCENTAGE RATE IN EACH CATEGORY OF THE INTERACTION PROFILE AND CUTTING POINTS ABOVE AND BELOW THE MEAN FOR THE DIAGNOSTIC REFERENCE POPULATION

Interaction-process analysis category	Percentage rates below the mean		Mean per- centage rate*	Percentage rates above the mean	
	1.96 SD	1.00 SD		1.00 SD	1.96 SD
1. Shows solidarity	.04	.74	2.97	5.10	8.50
2. Shows tension release	.24	2.40	8.17	13.90	22.20
3. Shows agreement	1.80	5.00	10.70	16.30	23.60
4. Gives suggestion	0	.77	6.56	12.40	22.20
5. Gives opinion	7.40	13.50	22.24	30.60	40.20
6. Gives information	9.90	17.90	28.72	39.60	51.10
7. Asks for information	1.20	2.90	5.88	8.70	12.50
8. Asks for opinion	.23	1.20	3.27	5.30	8.30
9. Asks for suggestion	0	.03	.60	1.10	2.20
10. Shows disagreement	.42	1.80	4.73	7.50	11.60
11. Shows tension	.01	.78	3.43	6.00	10.20
12. Shows antagonism	0	.21	2.41	4.40	8.20

* The means for each category are obtained directly by averaging the original percentage rates, rather than by retransformation back from the arc-sine means. Due to the nature of the transformation there are small losses in transforming back. Rather than leave the user with the uncomfortable feeling that there is some mistake in the table, we have avoided the transformation for the mean values shown.

What the user of Table 2 wants, we shall assume, is to be able to take a raw profile in which he has a special interest, to standardize it by computing percentage rates and (by comparing his profile with the mean rates and the cutting points on Table 2) to locate those particular rates on it that are unusual enough to merit special explanation. Then he may want to refer to Table 1 to see in what kind of condition unusual rates of this kind have been observed before. The choice of cutting points above and below the mean for locating "unusual" rates in this sense is essentially a pragmatic matter. From our experience, we believe that points one standard deviation above and below the mean on each category may prove generally useful for diagnostic inspection and clinical interpretation of single profile.

D. PROFILES FROM PARTICULAR STUDIES

We shall now examine briefly, one by one, the profiles in the reference population and the conditions under which they were obtained. This review will allow the reader to form an impression of the usefulness of the cutting points suggested and will provide him with an annotated bibliography of the studies from which he can obtain a more complete description of the conditions. If most of the rates appearing outside the suggested cutting points prove to be plausibly related to major features of the situation, task, type or condition of subjects as described by the investigator, we shall believe that the points have been well chosen.

1. *Bales' Case-Discussion Problem*

A group of students, usually five in number, unknown to each other, obtained through the employment office or by writing to a random sample, meet in a laboratory room with microphones and one-way mirror which are explained to them as they enter (2, 3, 4, 19). They are not introduced to each other by name, and no leader is appointed. The task of each session is the discussion of a human-relations case, a five-page presentation of facts about a problem facing an administrator in his organization. Members are given copies of the case separately to read ahead of time and are told that although each is given accurate information the experimenter intends to leave them uncertain as to whether they each have exactly the same range of facts. After they have been read by the members individually, the cases are collected to prevent direct comparison of typed copies; although members are allowed to take notes. The task defined for each session is to assemble the information, to discuss why the people involved in the case were behaving as they did, and to decide what should be recommended as action for the solution

to the problem presented. The group is asked to time themselves to 40 minutes and to dictate the group solution for the sound record in the final one or two minutes of the meeting. Then the experimenter leaves the room. The group is not arbitrarily stopped at the 40 minute limit, but is allowed to take a little longer if necessary.

Inspection of the summary profile of many such groups [see Table 1] shows that in spite of the pains taken in the experimental situation to highlight the desirability of communicating factual information, the rates of asking for information and giving it are both below the cutting point; however, the rate of giving opinion is above the suggested limit, identifying this problem situation as one with a notable requirement for the communication of inferences and value judgments. Rates of showing both agreement and disagreement are also high. (By "high" or "low" the authors mean above or below the suggested arbitrary cutting point one standard deviation from the mean.) The high rates of showing agreement and disagreement are plausibly related to the presumed status equality of members (no role differences established and symmetrical distribution of information) and to the requirement for a concrete and specific group decision. Under these conditions, most problem-solving attempts by a given member tend to obtain a reaction from some other, whether positive or negative.

2. Slater's Post-Task Social Session

The situation is that described earlier, just after the group members have individually completed a short questionnaire asking them to evaluate and describe each other on several different dimensions and to indicate feelings of liking and disliking for each other and the task situation (20). The experimenter says that he has one additional short questionnaire, but that he would like the members to spend a few minutes talking informally before he gives it. He says, "just talk about anything you like, I'll be back in about 10 minutes." He then leaves, and stays away from 13 to 15 minutes.

The summary profile shows that under these conditions the rates of asking for suggestions and giving suggestions drop to zero. The showing of antagonism also disappears. The rate of asking for information rises above the cutting point as the members "make social conversation" by asking about names, home towns, mutual acquaintances, courses, places of residence, and other mutual topics of reference. The rates of showing solidarity and tension release are also high, as cigarettes are passed around, and smiles, jokes, and laughs are exchanged.

3. *Couch's Prediction and Self-Revelation's Task*

The groups are essentially like those described earlier, but each group is now in a third session of the same set of members, with an expectation of three more sessions to follow (7). Each member answers privately a single item from a personality test or a values questionnaire. He is then asked to predict or guess how each of the other members in the group will answer the same item and to record privately his prediction. Members are then asked to reveal their answers to each other, note their errors, and discuss with each other in any way they like. Their aim is to improve their predictions on the next item. They go through a number of such cycles.

The rate of showing tension is high under this set of conditions; also, the rate of showing solidarity. This is presumably a stressful situation for several reasons. Not only is there a requirement to reveal to other group members feelings and characteristics that might otherwise be kept private; but success in prediction is difficult, failure is immediately revealed, and in the process of failing one may inadvertently reveal misconceptions he has of the other that may not be complimentary. Presumably the high rate of showing solidarity reflects an attempt to keep relationships from deteriorating in the face of these threats.

4. *Borgatta's Three-Man Groups Planning Role Playing*

One hundred twenty-six enlisted men at an Air Force Base are assigned to temporary duty to participate in an experiment (5). They are recruited from different organizations on the base, to minimize acquaintance. They meet in a new observation room with a one-way mirror. The experimenter explains that the purpose of the experiment is to observe how small groups work together. He says observations will be made of role-playing sessions which they themselves plan. Each man participates in four sessions of 48 minutes each with two new participants each time. About half of each session is given to getting acquainted and planning the role-playing situation; the other half consists of the role playing. The planning part of the session is observed, as well as the role playing proper. The profile we show is that for the planning half of the session.

Under these conditions only two rates are different enough to focus on as a diagnostic problem. Showing disagreement and showing antagonism are below the cutting point. The situation was not designed as it was in order to depress these types of interaction; however, it is not unreasonable to suppose that enlisted men ordered on temporary duty and placed with others

they do not know under observation in a room with a one-way mirror might imagine their superior officers to be behind the mirror evaluating their performance, and might tend to be "on their best behavior." It may be noted that while the rate of showing tension is not above the cutting point, it is markedly higher than the rate of showing disagreement. Our experience has led us to believe that this inversion is a presumptive indicator of stress due to inhibition of negative reactions. The two preceding profiles both show this pattern also.

5. *Mills' Three-Man Groups Constructing Projective Stories*

Three students, obtained through the employment office, gather in the observation room and are briefly told about the microphones and the mirror (14). The experimenter gives each of them a picture, one from the standard set originated by Murray for the Thematic Apperception Test, and allows a short time for each to look at his own picture separately. The pictures are then collected, and the experimenter asks the group to construct a dramatic story that in some way weaves the three pictures together in an account of what has gone on before, what is going on now, and what the outcome will be. He then leaves the group. Other groups of three see all three pictures jointly instead of separately, thus changing to some extent the problem of assembling information. The profile we show sums scores from both of these two sets of conditions.

The profile under these conditions shows a low rate of giving information. In spite of the fact that under one variation each subject must describe his picture to the others, the rate is low. One of the intended characteristics of the pictures, of course, is that they shall contain a minimum of clear-cut factual information. They are vague and ambiguous. It is difficult to make statements about them without inference, evaluation, and other "projected" elements from the subject's own experience. On the other hand, the "opinions" produced are not clearly differentiated from suggestions because the task involves an accretion of such elements only loosely strung together into a dramatic form rather than a tightly organized structure of carefully selected elements in a highly reasonable arrangement. Almost anything said may be taken as a suggestion by the other. Under these conditions the rates of asking for suggestion and giving suggestion are high and so is the rate of showing disagreement. The mood, presumably, is one of phantasy comparatively unchecked by fears of offending the other members or violating logical demands. There is no definite feedback from the environment that would indicate one suggestion is much better than another.

6. *Olmsted's Four-Man Groups Playing a Board Game and Making up a Story*

Twenty-four groups of four men each are composed of volunteers from a university population (15). Each group meets for two 20-minute sessions. During the first session the group members play a board game with the experimenter as an opponent. During the second session the members rank a series of paintings in order of group preference and then make up a story based on the paintings. Before the sessions begin the group members are given special instructions concerning the best way to go about solving the problems. Half of the groups are told that it is best to pay particular attention to members' feelings, while the other half are told it is best to be business-like and not to get involved with personalities.

Since one of the tasks (making up a story about a set of pictures) is similar to that used by Mills it is interesting to note that rates of giving and asking for suggestion are high and giving information is low. However, Olmsted's groups are high on disagreement (although low on antagonism); whereas Mills' groups were high on agreement. This difference may be connected with the need for greater logical rigor in the board game as compared to that in the more projective task. The board game (like the projective task), however, provides visually complete information to all members; thus reducing the need for verbal communication of information. Also, the two tasks are similar in requiring a large series or collection of specific suggestions for communication to the experimenter.

7. *Slater's et al.'s Four-Man Groups Under the Influence of Lysergic Acid*

Twenty-four male subjects, all college students, are formed into six four-man groups (21). Each group meets twice, once in a control meeting under normal conditions and once for a period of one and one-half hours following the administration of lysergic acid diethylamide (LSD). Three of the groups receive LSD prior to their first meeting; three, prior to their second meeting. For each meeting the subjects are brought into a small room and told to sit around a rectangular table. An observer with an interaction recorder is seated at another table opposite the subjects. The subjects are then instructed to read a case study of a human-relations problem in an administrative situation, discuss it, and within 40 minutes reach a group decision as to how the problem should be handled (Bales' standard task). The profile given in Table 1 represents the average interaction pattern for the six groups under the influence of LSD.

In these groups the rates of showing of solidarity and tension release are extremely high. Both percentage values are more than 1.96 standard deviations above the mean. Slater reports that even subjects with depressed or withdrawn reactions participate in this apparent hilarity, however involuntary their participation may seem. The task-oriented parts of the profile are similar to those of other groups on the Standard Case Discussion, except for the rate of giving opinion, which is less than half as high as for normal groups on this task. The additional fact that asking for opinion is below the cutting point seems to indicate a marked drop in concern for logical analysis or consensual evaluation in these groups. In general the subjects seem to be using the group situation to relieve individual tensions rather than to meet the demands of the task.

8. *Lennard's et al.'s Subjects Under the Influence of Lysergic Acid*

The subjects in this experiment are four college graduates, three female and one male, who were considered nonpsychotic on the basis of psychiatric interviews and a battery of clinical psychological tests (13). The subjects are paid to participate in a discussion of "The place of women in society." The discussion is recorded, transcribed, and scored from the typescript. All of the subjects participated under the influence of lysergic acid. Two of the subjects received 50 micrograms of LSD 25 before the session; the other two, 100 micrograms of LSD 25.

In comparing the interaction profile of this group under the influence of LSD with the same group in normal conditions, Lennard *et al.* note that verbal output is restricted and that there is a marked reduction in negative interpersonal responses. Compared with other groups in our reference population this group is not markedly low on negative responses, though it is below the mean; however, the rates of giving and asking for opinion are high. The authors attribute this rise to the attempts of the members to restore cognitive clarity despite their felt impairment due to the LSD. These rates are in marked contrast to the rates obtained by Slater in the study just described. Lennard's groups do not have high rates of showing solidarity and tension release. It appears that the reaction in Lennard's groups is quite different from that in Slater's. It is possible that the topic and the sex composition of the group constituted a basically different set of conditions for this group—a topic that could not be ignored. It is also possible that acts of tension release were not as evident in the typescript as they might have been in live recording.

9. *Takala's et al.'s Groups Drinking Brandy and Beer*

The subjects are employees of the Finnish government chosen from volun-

teers in consultation with the company's doctor and the supervisor concerned (25). They meet in eight groups of four to six members immediately after work in a room suitably prepared with tables and tablecloths. Each group discusses a series of three topics during a five-hour testing period. A 15-to-20-minute sample of interaction is observed for each topic. During one session brandy is served; during another, beer. The profile in Table 1 is a combination of the observations under both the brandy and beer conditions.

The authors note that their Finnish groups not under the influence of alcohol are different from ordinary American laboratory groups. The low frequencies of showing solidarity and tension and asking for information and opinion may be attributable to cultural differences rather than intoxication. The rate of giving opinion, which is high according to our standards, is nevertheless lower than in the Finnish normal groups; thus the effect of alcohol in these groups is to depress the rates of activity in the task categories and increase the amount of disagreement and antagonism.

Perhaps the most salient aspects of this profile are the high rates of showing disagreement and antagonism and the low rate of showing tension. These are effects one would plausibly expect from the action of alcohol. The low rate of showing solidarity may not be characteristic of the action of alcohol under other conditions. It would be worth testing whether or not a more usual pattern is an early increase in solidarity later followed by an increase in antagonism. The study showed that the groups drinking beer tended to be more positive than those drinking brandy.

10. *Psathas' Therapy Groups*

Two therapy groups, each with four patients, were seen by the same therapist for one year in psychoanalytically oriented psychotherapy (16). Nine meetings of the two groups are selected for analysis: three at the beginning of therapy, three at the middle, and three at the end. All the patients have psychosomatic problems, predominantly asthma. The groups meet in a room wired for sound recording. An observer behind a one-way mirror records who is speaking to whom. No detailed notes are made of the content of what is being said except to note the first and the last words of each person's burst of speech. However, nonverbal acts (such as nods, gestures, biting fingernails, and looking away) are scored according to the Bales' category system. The verbal interaction is scored by playing back a tape recording of the meeting. To this record is added in the proper places the nonverbal acts that have been directly observed.

In the profiles for these therapy groups both giving suggestions and asking

for suggestions are low, an indication that the members are not concerned with immediate and specific solutions. In the permissive atmosphere of these therapy groups it is not necessary for the group members to reach consensus on the solution to common problems.

As an explanation of the high rate of showing agreement, Psathas (16) suggests that an

... important reason for the maintenance of a high level of positive acts throughout all phases for therapy groups is probably the continued need to encourage patients to present their problems. The therapist is often high in simple acts of agreement and probably many of these acts are nods of the head or simple phrases indicating that he is listening and is interested in having the patient go on. Such acts, which are also initiated by patients, function to facilitate the presentation of material for discussion.

11. *Winter's Prison Therapy Groups*

Four therapy groups of 10 to 12 men each are observed during their usual one-hour therapy sessions in a state prison (26). The observer, who is seated in the back of the room, later checks his scores by listening to a tape recording of the session.

The therapist follows a method of guided group interaction which emphasizes the role of the leader in the therapy process. The leader's initiative is evident in the high rate of asking for opinion since the leader's interaction accounted for about half of the activity in this category for each group.

12. *Roberts' and Strodtbeck's Groups of Paranoid-Schizophrenic and Depressed Patients*

Two groups of five male patients, one a group of paranoid schizophrenics and the other a group of depressed patients, are observed as they reach a group decision under the leadership of a psychiatrist (18). During the discussion each of the observers notes the target of one subject's action and scores all the nonverbal acts that could not be recovered later from a tape recording. Later, when the tape recording of each session is transcribed, the observers' notes are added to the final protocol. The interaction-process scoring is then done from the transcribed protocols by persons who observed the sessions and were trained in the Bales' technique.

In the combined group profile shown, two categories of acts differ from the average. Showing antagonism is high, primarily a reflection of the generally "hostile" characteristics of the paranoid-schizophrenic patients. The high rate of showing tension reflects the high-tension level of both types of patients.

13. *Evans' Therapy Groups*

Two therapy groups of 10 persons each are composed of male, adult patients at a psychiatric clinic for veterans (8). Each group meets for 16 sessions. Two observers score each session through a one-way mirror using an earlier version of the Bales' category system which contained 49 categories. Acts on which the observers show marked differences in scoring are omitted, and the remaining categories are combined into 12.

Both patients and therapists are high on giving information and low on asking for information. Most of the interaction of the patients in this type of therapy appeared to be giving information about the self. The high rates of asking for opinion and suggestion are primarily the result of the activity of the therapists as they sought to draw out the patients.

14. *Hare's et al.'s Psychiatric-Stress Interviews*

A series of 24 interviews with adults are observed through a one-way mirror in an initial psychiatric interview (11). The interviewer follows a partially standardized pattern, based on Chapple's methods. The same experienced psychiatrist interviews all 24 patients. The majority of the subjects are outpatients in a large New England hospital. They are predominantly lower class with an average age of 36. Seventeen are female; seven, male.

The interviewer has certain rules to follow during each of the five subperiods of the interview. The rules govern the timing of his responses, but not the content of what he says. Periods 1, 3, and 5 are "free" in that no behavior stress is applied. For each patient utterance the doctor must respond within one second after the patient has finished speaking. In addition, the interviewer must speak for approximately five seconds during each of these utterances. In Period 2, a silence stress is applied to the subject as the interviewer now "fails to respond" by waiting up to 15 seconds after the subject has finished speaking. If the subject still has not spoken after the 15 seconds, the interviewer speaks again for five seconds. This period ends after 12 "failures to respond" or 15 minutes, whichever is shorter. Period 4 is one of an interrupting stress. Each time the subject speaks the interviewer waits for about three seconds and then interrupts with a five-second utterance for a total of 12 times. If the subject continues to speak through this interruption, the interviewer does not interrupt again until the subject has finished his utterance. The time limit for Period 1 is 10 minutes; for Periods 3 and 5, five minutes. In the average interview, Period 2 is eight minutes long and Period 4 is four minutes, bringing the average total interview to about 32 minutes.

The interaction pattern that results from this combination of predetermined formal interaction pattern and a process devoted to fact gathering rather than group problem solving differs markedly from the average pattern of the 21 types of groups. The doctor plays the major part in asking for information to which the patient responds by giving information. In addition to the generally high level of tension shown by the patient throughout the interview, the doctor also contributes to the tension score as he fails to respond during the silence stress. No suggestions are asked for by either party, and only a few direct suggestions are made by the doctor. Low rates of joking and laughing and showing solidarity indicate that the tension throughout the interview is unrelieved. Low rates of agreement and disagreement reflect the fact that group consensus is not required by the task. The patient's acts contribute most heavily to the total interaction profile since he contributes approximately four acts for every act of the doctor.

15. *Strodtbeck's and Mann's Jury Deliberations*

The participants in these mock-jury deliberations are drawn from the regular jury pools of the Chicago and St. Louis courts (24). The jurors listen to a recorded trial. They then deliberate and return their verdict—under the customary discipline of bailiffs of the court. The deliberations are recorded with two microphones to facilitate binaural identification of individual participants. The recordings are fully transcribed, and these protocols are in turn scored in terms of interaction-process categories. The scoring is done by an assistant who listens again to the recording and has available the indications of nonverbal gestures made by the original observer. The level of interscorer reliability is checked before the scoring begins and rechecked periodically while scoring is in process.

The 12 protocols utilized in computing the summary profile shown are the final 12 of a set of 30 in which jurors considered an auto negligence case. Seventeen of 144 jurors originated less than five acts each and have been dropped from the tabulations.

In the task area this summary profile does not differ from the average of the 21 profiles. However, the jurors show more agreement and less tension and antagonism than the average group. A higher-than-average amount of agreement is also found in other types of discussion groups in which the group members are required to reach consensus on an issue [see Table 1].

16. *Strodtbeck's Husbands and Wives*

Each of 10 couples from three cultures, Navaho, Texan, and Mormon, are asked to discuss the differences in their ratings of some mutual friends (23).

Each couple is asked to pick three reference families with whom they are well acquainted. The husband and wife are then separated and requested to designate which of the three families appears most satisfactory with respect to each of a series of 26 conditions, such as "Which family has the happiest children?" "Which family is the most religious?" "Which family is the most ambitious?" After both husband and wife have individually marked their choices, they are requested to reconcile their differences and indicate a final "best" choice from the standpoint of their family. The discussions are recorded in the field by portable sound equipment powered from a truck. The recordings are then transcribed and, in the case of the Navaho, translated into English before being scored.

Although Strodtbeck found differences in interaction profiles between cultures and between the most-talking and least-talking spouse in each culture, the only category in which the summary profile for all 30 couples is distinguished from the others in Table 1 is a high rate of agreement. This seems to reflect the fact that the task called for reaching agreement on a number of judgments.

17. *Quade's College Couples*

Sixty college couples are observed from behind a one-way mirror, while discussing two problems for an average time of 31 minutes per problem (17). Since couples vary considerably in the time it takes to solve the problem (range: 10 to 65 minutes) the scores for the first 15 minutes of each session are used. In those cases in which the couple talks for less than 15 minutes an adjustment is made to yield comparable data.

For the discussion of the recreation problem, given first, a large calendar is placed on a work table. The couple is instructed to go through the calendar day by day and construct a schedule of all their leisure and recreational activities. Each makes a separate schedule. One spouse is not allowed to enter an activity on his schedule without the consent or agreement of the other.

For the budget problem, the couple is told:

Please construct a budget of all your expenditures for a typical month. The amount you are to use in building your budget is your own present total take home pay plus 25 per cent. You cannot spend more than one quarter of the 25 per cent on any one thing. Keep separate records of your budget. Do not enter anything without the consent or agreement of your spouse. When you have finished the budgets will be identical.

After presenting each problem the observer leaves the room where the couple is working and records their interaction through a one-way mirror.

Operating under these task instructions, which require immediate and specific consensus at many points throughout the total problem, the rates of giving and asking for suggestions are higher than average. The low rate of showing solidarity probably reflects several factors. First, as Quade notes in his analysis of the data, married couples devote a very small percentage of all their actions to building solidarity, presumably because they are more secure about the level of solidarity they have obtained. Second, half of the couples in Quade's sample were classified as maladjusted; and these couples tended to have even lower rates of showing solidarity. Finally, according to Quade, the recreational problem involved less potential conflict than the budget problem and therefore required fewer attempts to restore equilibrium by showing solidarity.

18. *Landsberger's Labor-Management Mediation Sessions*

Twelve mediation cases are analyzed from transcribed recordings (12). Each of the groups consists of employees and officials of labor unions; of employers and their lawyers or other adjunct personnel; and of a labor mediator. It is the mediator's function to assist the two parties in reaching an agreement after bipartite negotiations between them have previously broken down. The cases are typical of those involving enterprises with less than 25 employees. Each case was settled or adjourned *sine die* at the end of one meeting—the one recorded. Each mediator has an office in which the mediation takes place. He differs from an arbitrator who typically hears a case, much as does a judge, and thereafter makes a formal award. The mediator does not make an award; he clarifies, persuades, and suggests.

The emotional involvement of the parties in the discussion is reflected in the higher-than-average amount of showing disagreement and antagonism. The amount of hostility within a session is negatively correlated with the success of the outcome of the session. The serious intent of the sessions is suggested by the low rate of showing tension release.

19. *Hare's Boy Leaders*

Twelve third-grade boys who have been identified by their teachers as leaders on the school playground are observed at play with their friends in their home neighborhoods after school (10). Six of the boys have been selected because they are group oriented; the other six, because they are self oriented. Two observers score interaction in the field. Each observer scores only one boy at a time, recording all of his acts directed to others and all acts directed to him by other group members. When possible, each observer makes half of the observations on each subject.

The average observation period in the neighborhood is four sessions for a total of 75 minutes (218 acts). Interaction with parents or observers is not recorded; only the interaction with other children is included in the profiles. Nonverbal interaction called for by the rules of the game is also omitted.

This average profile for boys' leaders and their peers is similar in several respects to that of Gruber's children playing with dolls. Both show a marked contrast with the Harvard-laboratory discussion groups. For the children, agreement and giving opinion are low; asking for information is high. The tendencies are just the opposite of those in the problem-solving groups. Rather than to deliberate and attempt to reach consensus on issues, these boy leaders and their followers interact very directly, with high rates of giving suggestions and showing antagonism.

20. *Gruber's Children Playing with Dolls*

The behavior of 13 preschool children judged to be high on Task orientation and 13 judged to be low on Task orientation is recorded as they play with a doll house and dolls in the presence of a psychologist (9). When each child is taken into the playroom he sees a doll house without dolls. He is then given dolls as he asks for them. As the child plays with the dolls for 40 minutes the psychologist takes longhand notes of the child's behavior and interaction. The protocols are scored later, using the Bales' categories.

Since the psychologist is primarily interested in how much task behavior the child will initiate if left to his own devices there is relatively little interaction with the experimenter. The intention of the experimenter is reflected in his low frequencies of agreement, giving opinion, and asking for opinion and suggestion. The high rate of tension represents the child's attempts to withdraw from play with the doll house. The high rate of showing antagonism, which was also evident for the boy leaders, would seem to indicate that children are freer in expressing hostility than adults except when the adults are intoxicated or under some emotional stress.

21. *Borgatta's et al.'s Students of Interpersonal Relations*

A class of 47 graduate students in a course on interpersonal relations is divided into five subgroups that meet for a two-hour discussion period each week (6). The class is made up of relatively mature persons (average age 34 years) representing a number of socially oriented professions. There is a slightly larger number of females than males. At the time the study takes place the groups have met for at least nine weeks.

The groups have little or no external supervision. For each session, persons are selected on an informal rotational basis to serve in the positions of leader, observer, and recorder. The latter two seem to be similar in form because they both require the incumbent to record what is going on, although from different points of view. There is an emphasis throughout the discussions on the analysis of the group's own process and on aspects of behavior that tend to reflect "democratic leadership."

Each of the five groups is observed twice, for approximately two hours each session. Although the investigator is in the room, he is excluded by agreement from the attention of the group. Remarks made during the sessions and in discussion after the second session indicate that individual members occasionally were conscious of the observer's presence; however, in most cases the groups operated in what they described as their normal manner.

These discussion groups and the husbands and wives observed by Strodtbeck are closer to the mean profile for all groups than any other groups in the sample. Both differ from the average by showing a rate of agreement that is one standard deviation above the mean. In this case the high amount of agreement seems to be a reflection of the emphasis on "democratic process."

E. SUMMARY

A number of investigators have used the same set of categories for interaction-process analysis. Twenty-one profiles from various studies are presented as a reference population. Each profile is either a summary for a given study or for a series of studies under the same conditions. An interaction profile for a group or some set of individuals can now be compared for diagnostic purposes with this reference population.

A profile from a given study is made comparable with others by transforming the number of scores in each category into a rate that is computed as the percentage of the total number of acts in all 12 categories. Means and standard deviations for the rates in each of the 12 categories were computed for the 21 reference profiles. Cutting points of one standard deviation above and below the mean appear to be useful for diagnostic inspection and clinical evaluation of single profiles.

REFERENCES

1. BALES, R. F. *Interaction Process Analysis: A Method for the Study of Small Groups*. Cambridge, Mass.: Addison-Wesley, 1950.
2. BALES, R. F., & SLATER, P. E. Role differentiation in small decision-making groups. In T. Parsons, R. F. Bales, *et al.* (Eds.), *The Family, Socialization, and Interaction Process*. Glencoe, Ill.: Free Press, 1955. Pp. 259-306.

3. ———. Social science groups. Unpublished data, 1955.
4. BERGER, J. Relations between performance, rewards and action-opportunities in small groups. Unpublished Doctoral dissertation, Harvard University, Cambridge, Mass., 1957.
5. BORGATTA, E. F., & BALES, R. F. Task and accumulation of experience as factors in the interaction of small groups. *Sociometry*, 1953, **16**, 239-252.
6. BORGATTA, E. F., COTTRELL, L. S., JR., & MANN, J. H. The spectrum of individual interaction characteristics: An inter-dimensional analysis. *Psychol. Reports*, 1958, **4**, 279-319.
7. COUCH, A. S. Psychological determinants of interpersonal behavior. Unpublished Doctoral dissertation, Harvard University, Cambridge, Mass., 1960.
8. EVANS, J. T. Objective measurement of the therapeutic group process. Unpublished Doctoral dissertation, Harvard University, Cambridge, Mass., 1950.
9. GRUBER, S. Task-orientation in the preschool child and its implications for early school adjustment. Unpublished Doctoral dissertation, Harvard University, Cambridge, Mass., 1952.
10. HARE, A. P. Situational differences in leader behavior. *J. Abn. & Soc. Psychol.*, 1957, **55**, 132-135.
11. HARE, A. P., WAXLER, N., SASLOW, G., & MATARAZZO, J. D. Simultaneous recording of Bales and Chapple interaction measures during initial psychiatric interviews. *J. Consult. Psychol.*, 1960, **24**, 193.
12. LANDBERGER, H. A. Interaction process analysis of professional behavior: A study of labor mediators in twelve labor-management disputes. *Amer. Sociol. Rev.*, 1955, **20**, 566-575.
13. LENNARD, H., JARVIK, M. E., & ABRAMSON, H. A. Lysergic acid diethylamide (LSD-25): XII. A preliminary statement of its effects upon interpersonal communication. *J. of Psychol.*, 1956, **41**, 185-198.
14. MILLS, T. M. A method of content analysis for the study of small groups. Unpublished Doctoral dissertation, Harvard University, Cambridge, Mass., 1951.
15. OLMSTED, M. S. Small group interaction as a function of group norms. Unpublished Doctoral dissertation, Harvard University, Cambridge, Mass., 1952.
16. PSATHAS, G. Phase movement and equilibrium in interaction process in psychotherapy groups. *Sociometry*, 1960, **23**, 177-194.
17. QUADE, A. E. The relationship between marital adjustment and certain interactional patterns in problem solving situations. Unpublished Doctoral dissertation, Ohio State University, Columbus, 1955.
18. ROBERTS, B. H., & STRODTBECK, F. L. Interaction process differences between groups of paranoid schizophrenic and depressed patients. *Internat. J. Group Psychother.*, 1953, **3**, 29-41.
19. SALISBURY, M. R. Judgments of personality in decision-making groups. Unpublished Doctoral dissertation, Radcliffe College, Cambridge, Mass., 1958.
20. SLATER, P. E., & BALES, R. F. Experimental groups—social session. Unpublished data, 1957.
21. SLATER, P. E., MORIMOTO, K., & HYDE, R. W. Social interaction in experimentally induced psychotic-like states. Paper read at American Sociological Society Meetings, Seattle, Washington, 1958.
22. SNEDECOR, G. W. Statistical Methods. Ames, Iowa: Iowa State College Press, 1937. Pp. 449-450.
23. STRODTBECK, F. L. A study of husband-wife interaction in three cultures. Unpublished Doctoral dissertation, Harvard University, Cambridge, Mass., 1950.
24. STRODTBECK, F. L., & MANN, R. D. Sex role differentiation in jury deliberations. *Sociometry*, 1956, **19**, 3-11.

25. TAKALA, M., PIHKANEN, T. A., & MARKKANEN, T. The effects of distilled and brewed beverages: A physiological, neurological, and psychological study. *Finnish Foundation for Alcohol Stud.*, 1957, 4.
26. WINTER, G. Value-orientations as factors in the organization of small groups. Unpublished Doctoral dissertation, Harvard University, Cambridge, Mass., 1952.

*Laboratory of Social Relations
Harvard University
Cambridge 38, Massachusetts*

*Department of Sociology
Haverford College
Haverford, Pennsylvania*

CORRELATES OF CIVIL-RIGHTS PARTICIPATION* 1. 2

University of Rhode Island and Fisk University

LESTER CARR AND S. OLIVER ROBERTS

A. INTRODUCTION

Examination of the literature on social attitudes raises questions as to the degree of consistency between verbal attitude expression and overt action (4, 8, 14). In a comprehensive study, Pace (11) analyzed the replies of some 2500 university alumni to a 16-page questionnaire that included seven "Activity" scales and nine Likert-type "Opinion" scales. On the basis of the small correlations found between the attitude and activity measures, Pace declared, "Opinions are the verbalized expression of attitudes, opinions are not action."

Campbell (3) takes the position that much of the so-called inconsistency between attitudes and action is "pseudoinconsistency" and would disappear if more reliable and valid overt-behavior measures were used and a wider variety of behaviors was sampled. In line with this, the Rosenzweig Picture-Frustration test, a Likert-type scale, a Guttman scale, and the H Technique were used to predict the willingness of college students to sign a petition in favor of Fair Employment Practices Legislation (6). These four attitude-measuring devices were able to predict the criterion almost equally well, with biserial correlations of .54, .61, .58, and .59 respectively. However, the criterion measure employed is questionable because the signing of a petition may be more representative of verbal behavior rather than of complex social-action involvement.

Searles and Williams (13) administered questionnaires to 827 students at three Southern Negro colleges and investigated the relationship of demographic and attitudinal variables to civil-rights activities. When their data were divided into three groups according to the "degree" of respondent's participation, they found "fewer than [a] chance number of statistically significant differences among the three."

* Received in the Editorial Office, Provincetown, Massachusetts, on May 11, 1964. Copyright, 1965, by The Journal Press.

¹ Based on a paper entitled, "Social action as related to verbalized attitudes, self-concept and background factors," presented at the 1963 American Psychological Association convention.

² The authors wish to acknowledge with deep appreciation the counsel of Dr. J. C. Nunnally, chairman of the department of psychology, and the psychology faculty of Vanderbilt University during the course of this study.

Recent activities of Negro American college students in "sit ins" and similar social-action participation provide unique opportunities to check expressed attitudes against overt behavior: i.e., to obtain measurable behavioral correlates of verbalized attitudes. In this study, three conventional attitude scales designed to measure attitudes toward social action were related to the level of involvement in and frequency of social-action activities; also, correlations among the various attitude and behavioral measures were compared. In addition to the primary methodological purpose, another aim was to explore some self-concept and background variables and their relationship to social-action behavior.

B. METHOD

1. Subjects

The total pool for the study consisted of 501 Negro American college students, freshmen and sophomores, in a predominantly Negro, private liberal-arts university located in the South. A group of 332 Ss (104 males and 224 females) was randomly drawn from the total pool and given the final form of the attitude scales.

2. Behavioral Measures

On the basis of student statements, social-action behavior was divided into the three levels as follows:

a. *Level 1 (actively involved)*. Participation in a minimum of one "sit in" or "stand in."

b. *Level 2 (moderately involved)*. Participation in some civil-rights activity, such as steering committees, making posters, etc.; but "no" participation in any "sit ins" or "stand ins."

c. *Level 3 (uninvolved)*. No participation of any kind in civil-rights activities.

Assignment of *S* to one of the levels depended on his satisfying both of the following criteria: (a) being identified by a student leader as having participated in a particular activity and (b) indicating on a questionnaire, civil-rights involvement. In addition, the total number of times *S* had been involved in civil-rights activities of any type was determined.

3. Attitude Scales

Three scales were employed: social distance, social action, and semantic differential. The Bogardus social-distance scale concerns "direct actions" rather than more general feelings about people or situations and seems especially appropriate for measuring attitudes toward "social action." The traditional

Bogardus groupings were applied separately to categories, such as Mexican, Puerto-Rican, lower-class Negro, Italian, Irish, etc. (2). An overall index of social distance was obtained by computing the sum of the groupings checked for a particular social category and adding all 16 social-category scores; therefore, a high score signifies less social distance: i.e., more social contacts desired with the various categories.

Nunnally (9), utilizing a semantic-differential type questionnaire, found that "understandability" was an important component of public reaction to the mentally ill, and "nervousness or anxiety" is the cardinal sign of mental disorder in public thinking. Using Nunnally's findings as a reference point, items concerning evaluation, understandability, and adjustment were designed for the present research (21 scales in all, randomly polarized):

Evaluation was represented by the following eight scales: kind—cruel, effective—ineffective, pleasant—unpleasant, intelligent—ignorant, good—bad, fair—unfair, wise—foolish, valuable—worthless.

Understandability was represented by the following six scales: predictable—unpredictable, familiar—strange, understandable—mysterious, simple—complicated, confusing—clear, straight—twisted.

Adjustment was represented by the following seven scales: relaxed—tense, stable—unstable, mature—immature, happy—sad, calm—anxious, unemotional—emotional, well-adjusted—maladjusted.

Nine groups considered as "concepts" were to be rated on a nine-point continuum on the aforementioned scales. These groups included Northern Negroes, Southern Negroes, Northern Whites, Southern Whites, Negro College Students Who Take Part In Civil-Rights Demonstrations, Negro College Students Who Do Not Take Part In Civil-Rights Demonstrations, White College Students Who Help Negroes In Civil-Rights Demonstrations, White College Students Who Do Not Take Part In Civil-Rights Demonstrations, and Myself. A score for each of the groups of scales (evaluation, understandability, and adjustment) was obtained by adding the separate ratings on the various scales composing a particular group. The higher the rating the more positive it was and indicated a higher evaluation or greater understanding or adjustment attributed to a particular concept.

Student remarks served as a basis for composing social-action statements. These statements were included in a Likert-type instrument that went through several stages of purification: an item-analysis was performed as described in Edwards (5, Chapter 6). Illustrative items are "The only way to make society better is to go out and change it" and "It is best to avoid people who are known in the community as outspoken supporters of civil rights."

4. *Background Factors*

A questionnaire was used for obtaining all background information and included the following variables: socioeconomic status (Minnesota Occupational Scale), region, family disorganization (married, separated, divorced), and degree of integration of elementary and high schools attended.

5. *Identical Concepts On Different Instruments*

In order to determine the extent to which the type of instrument employed affects expressed attitudes, randomly selected sections of freshmen and sophomores (80 Ss from the original pool) were given Bogardus-type and semantic-differential instruments—both of which included the same two terms (Northern Whites and Southern Whites).

C. RESULTS

1. *Instrument Reliabilities*

The following four-month test-retest reliability coefficients were found: social-distance scale, .93; social-action scale, .80; and for all nine concepts composing the semantic-differential scale, the median correlation for each of the groups of items was .45 for evaluation, .43 for understandability, and .43 for adjustment [see (10), Chap. 3]. The highest semantic-differential reliabilities were found for the "Myself" concept (evaluation .57, understandability .53, and adjustment .62).

2. *Attitudes Versus Behavior*

Examination of relationships among all attitude variables and the behavioral measures reveal no correlation coefficient that accounts for more than nine per cent of the variance in social-action behavior. In fact, 80 per cent of the Pearson product-moment coefficients are below .15; therefore, if there is a relationship between verbalized attitudes and actions (as defined in this study) it must be small indeed.

3. *Attitude Scales and Relationships*

When the three attitude instruments (each including several measures of attitudes toward social action) were related to each other no coefficient greater than .29 could be found (7). [See also Table 1.] As demonstrated in

TABLE 1
CORRELATIONS BETWEEN MEASURES OF SOCIAL-ACTION INVOLVEMENT AND ATTITUDES TOWARD SOCIAL ACTION, SOCIAL DISTANCE, NEGRO COLLEGE STUDENTS WHO TAKE PART IN CIVIL-RIGHTS DEMONSTRATIONS, AND SELF*

Variable	Sex	1	2	3	4	Variable	5	6	7	8	9	10
1. Social-action level	M											
	F											
2. Social-action frequency	M	.64										
	F	.39										
3. Social-action Scale: total score	M	.26	.29									
	F	.25	.10									
4. Social-Distance Scale: total score	M	.14	.14	.29								
	F	.08	.03	.15								
Negro college students who take part in civil-rights demonstrations:												
5. Evaluation	M	.25	.25	.26	-.10							
	F	.15	.10	.08	-.13							
6. Understandability	M	.17	.15	.17	.13	.68						
	F	.15	.13	.14	-.08	.49						
7. Adjustment	M	.07	.05	.20	.07	.60	.69					
Myself	F	.04	-.01	.16	-.04	.43	.47					
8. Evaluation	M	-.09	-.02	-.15	-.06	.32	.28	.34				
	F	.03	.10	.02	-.09	.42	.25	.24				
9. Understandability	M	-.10	-.08	.01	.03	.27	.38	.27	.58			
	F	.10	.20	.01	-.12	.20	.27	.17	.46			
10. Adjustment	M	-.12	-.18	-.08	-.01	.27	.25	.25	.67	.68		
	F	.05	.06	.00	-.08	.27	.27	.39	.48	.44		
Instrument reliability												
			.80	.93	.41	.43	.41	.57	.53	.62		

* An r of .20 (males) and .14 (females) is required for significance at the .05 level. N for males = 104; N for females = 228. Higher social-action level and frequency are considered as positively related to the more favorable end of the other scales: for example, less social distance is favorable, and rating on a semantic-differential scale is scored towards the positive (favorable) adjective.

Table 2, small and nonsignificant correlations (.10 is the largest) were found between scores obtained from concepts and associated groups of items that are identical except for being embedded in different-type instruments.

TABLE 2
CORRELATIONS BETWEEN IDENTICAL CONCEPTS APPEARING IN
DIFFERENT INSTRUMENTS
($N = 80$)

Bogardus scale	Evaluation		Semantic-differential scale Understandability		Adjustment	
	Southern whites	Northern whites	Southern whites	Northern whites	Southern whites	Northern whites
Southern whites	.10	.02	.02	.05	.05	.03
Northern whites	.05	.07	.10	.04	.01	.02

4. Background Factors and Social Action

Table 3 shows small correlations between the behavioral measures and a variety of background factors.

TABLE 3
CORRELATIONS BETWEEN MEASURES OF SOCIAL-ACTION INVOLVEMENT
AND SELECTED BACKGROUND FACTORS*

Social action	Sex	North-south	Integration		Parents' relationship	Occupation	
			Elementary school	High school		Father	Mother
Level-action level	M	-.22	.18	.21	-.12	.03	-.07
	F	-.04	-.08	-.01	.02	.09	-.01
Frequency	M	-.26	.23	.18	-.12	.09	-.07
	F	-.09	.02	.07	.01	.16	-.04

* An r of .20 (males) and .14 (females) is required for significance at the .05 level. N for males = 104; N for females = 228. In the case of the background variables it is assumed that North, higher degree of integration (elementary or high school), parents together, and "higher" occupation of either father or mother are *positively* related to greater amount of social-action involvement.

D. DISCUSSION

The present study reveals no factors (attitudes or otherwise) significantly related to social-action behavior. These findings are in agreement with those reported by Roberts and Carr (12) in 1961. The small correlations are not due to insufficient reliability. Except for the semantic-differential scale, reliabilities are generally high for all measures; and even with those for the semantic-differential scale, the reliabilities are sufficient to support relatively large correlations. Certainly the number of cases was large enough to leave

little doubt of the inconsequentiality of the correlations. One possibility for finding low correlations is attenuation of unreliability; however, even after attenuation corrections were made relatively low (and insignificant) correlations are found. The attitudes are of opinions highly "crystallized": for example, the opportunity to involve one's self in civil-rights activities was clearly present over a period of two years. The university campus is located within walking distance of the demonstration areas, and participation was fully supported by both the university administration and faculty.

One possibility to account for the slight correlations is that individual characteristics play a relatively minor role; other, more primary determinants of social action being a wide range of social-situational or field forces. Evidence to support this position was provided by intensive open-end taped interviews with approximately 10 per cent of the *Ss* at each of the three social-action levels. For example, some reasons given by students at Level 3 (uninvolved) for not taking part in civil-rights activities included parental pressure, fear of names being printed in hometown newspapers, numerous complaints about the policies and actions of local and national civil-rights organizations, and poor grades; while some reasons given by those at Levels 1 (actively involved) and 2 (moderately involved) for participating included a social-activist's talk, newspaper editorials and general reports, television accounts of violence to Negro students, individual or group persuasion by fellow students, educational literature, and appeals of local and national civil-rights organizations.

Perhaps the small and nonsignificant correlations were determined in some unknown way by factors peculiar to the investigation. If not, the results point to a surprising inconsistency among three widely used methods of studying verbalized attitudes and suggest that none of these (as well as some widely used background and self-concept variables) is predictive of real-life behavior. In essence, it may well be that (at least in the case of Negro American college students) the crucial determinants of social-action involvement are immediate, contextual field conditions; and with the impact of these social-situational field forces, any consistent individual differences are soon overcome. Since for the majority of Negro students their personality is not static and can be viewed as an open system in the process of "becoming," this has implications for the present desegregation struggle (1). As our interview material emphasized, educational agencies, news media, and civil-rights groups can play a significant if not primary role in determining social attitudes and stimulating or inhibiting student involvement in social issues.

E. SUMMARY

The level of involvement in and the frequency of social-action activities were examined in relation to several conventional scales designed to measure attitudes toward social action, self-concept, and background factors. The Ss were 501 Negro American freshmen and sophomores in a predominantly Negro, private liberal-arts university located in the South. A group of 332 Ss (104 males and 224 females) was randomly drawn from the total pool and given the various scales. The remainder of Ss were utilized for instrument purification and pretesting. Product-moment correlations were calculated among the different attitude scales and behavioral measures.

Substantial test-retest reliability coefficients were found for the attitude instruments and high agreement was indicated for the behavioral measures. However, practically no significant correlations were found between any of the measures of expressed attitudes and overt behavior. Furthermore, small correlations were found between the several attitude scales supposedly measuring the same thing. Moreover, small correlations were found between responses given to identical concepts but imbedded in different-type instruments. Added to all this, examination of all correlations revealed no factor that accounted for more than nine per cent of the variance in social-action behavior. Reasons to account for these findings were discussed.

REFERENCES

1. ALLPORT, G. W. The open system in personality theory. *J. Abn. & Soc. Psychol.*, 1960, **61**, 301-311.
2. BOGARDUS, E. S. Measuring social distance. *J. Appl. Sociol.*, 1925, **9**, 299-308.
3. CAMPBELL, D. T. Social attitudes and other acquired behavioral dispositions. In S. Koch (Ed.), *Psychology: A Study of a Science*. Vol. 6 Investigations of Man as Socius: Their Place in Psychology and the Social Sciences. New York: McGraw-Hill, 1963.
4. COREY, S. N. Professed attitudes and actual behavior. *J. Educ. Psychol.*, 1937, **28**, 271-280.
5. EDWARDS, A. L. *Techniques of Attitude Scale Construction*. New York: Appleton, 1957. Chap. 6.
6. KAMENETZKY, G. C., BURGESS, G. G., & ROWEN, T. The relative effectiveness of four attitude assessment techniques in predicting a criterion. *Educ. & Psychol. Meas.*, 1956, **16**, 187-194.
7. KERR, W. A. Untangling the liberalism-conservatism continuum. *J. Soc. Psychol.*, 1952, **35**, 111-125.
8. LA PIERRE, R. T. Attitudes *vs.* actions. *Soc. Forces*, 1934, **13**, 230-237.
9. NUNNALLY, J. C. *Popular Conceptions of Mental Health: Their Development and Change*. New York: Holt, Rinehart, & Winston, 1961.
10. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957.

11. PACE, C. R. Opinion and action: A study in validity of attitude measurement. *Educ. & Psychol. Meas.*, 1950, **10**, 411-419.
12. ROBERTS, S. O., & CARR, L. "Social action" participation as related to selected variables for Negro American college students. *Amer. Psychologist*, 1961, **16**, 398.
13. SEARLES, R., & WILLIAMS, A. J., JR. Negro college students participation in sit-ins. *Soc. Forces*, 1962, 215-220.
14. THURSTONE, L. L., & CHAVE, E. J. *The Measurement of Attitude*. Chicago: Univ. Chicago Press, 1929.

*Department of Psychology
University of Rhode Island
Kingston, Rhode Island*

given a series of Thematic Apperception pictures, and protocols were obtained in the manner proscribed by McClelland (2). These protocols were scored for need achievement.

The paired-adjective continua employed were 12 in number and were arranged so that each pair was separated by a seven-point scale. Each subject was invited to rate his conception of "time" on each of the scales. It appeared on inspection that subjects varied considerably both in the means tendency on the scales and, more important, on their range of ratings. Accordingly, the raw ratings were not subjected to analysis, but were converted into normalized *T* scores for each subject; thus equating for each subject the mean and the standard deviation of his ratings on the 12 scales.

C. RESULTS

Each rating scale was correlated with every other and with the need-achievement measure described earlier, yielding a 13-by-13 matrix of correlations. Perhaps the first and most instructive thing to be noted is the correlation obtaining between the need-achievement measure and ratings on each of adjective continua. They are given in Table 1 in order of the correlation of the first adjective of the pair with need achievement.

TABLE 1
CORRELATIONS BETWEEN NEED ACHIEVEMENT AND ADJECTIVE CONTINUA

Continuum	<i>r</i>	Continuum	<i>r</i>
clear—vague	.24	tense—relaxed	.00
sharp—dull	.18	cold—hot	— .06
empty—full	.08	sad—happy	— .11
active—passive	.06	deep—shallow	— .11
young—old	.06	soothing—irritating	— .11
fast—slow	.02	small—large	— .20

Individually, none of the correlations reported attains the .01 level of confidence (i.e., .28), and only one exceeds the .05 level of confidence (i.e., .21); thus, it is clear that our results are not comparable in their significance with those obtained with our Metaphor Test. All the same, it may be instructive to note that the ordering of our adjective scales does not appear random. If we consider those adjectives that correlate at least one standard deviation from zero, we find that need achievement correlates positively with "clear," "sharp," "happy," "shallow," "irritating," and "large." It correlates negatively with "vague," "dull," "sad," "deep," "soothing," and "small." There appears a confirmation of that pattern manifest in the correlations of time metaphors with

need achievement: namely, the tendency for energetic and dynamic images of time to be associated with high need achievement.

Next, we proceeded to a factor analysis of our matrix, comprising correlations between our 12 adjectival continua and our measure of need achievement. Two factors were obtained, though for our purposes we shall present only the first and clearly most incisive factor. The loadings on this first factor are shown in Table 2.

TABLE 2
FIRST FACTOR LOADINGS EMERGING FROM FACTOR ANALYSIS OF 12 ADJECTIVE
CONTINUA AND NEED-ACHIEVEMENT MEASURE

Adjective pair	Loading	Adjective pair	Loading
active—passive	— .53	small—large	.10
sharp—dull	— .49	deep—shallow	.39
fast—slow	— .46	soothing—irritating	.45
tense—relaxed	— .37	empty—full	.48
clear—vague	— .37	sad—happy	.63
Need achievement	— .33	cold—hot	.64
young—old	— .24		

It will be appreciated that the loadings of each of these adjective scales may be reversed in sign by reversing the paired adjectives. If this is borne in mind, we find that need achievement is factorially proximate to the following: active, sharp, fast, tense, clear, hot, happy, full, irritating, and shallow. It is factorially alien, on the other hand, from the following adjectives: passive, dull, slow, relaxed, vague, deep, soothing, empty, sad, and cold.

D. DISCUSSION

The results that we report here are congruent with the implications of our earlier study relating preference for time metaphors to measured need achievement. In a sense, the present study may be thought of as a replication with a variation in one of the dimensions: namely, the mode of measured attitudes toward time. It is of some interest to us that we obtained more incisive results from our metaphor scale for time than from the semantic differential. Both, however, have indicated that high need achievement is associated with an energized and dynamic sense of the flow of time; while low need achievement is associated with passive and static attributes. Elsewhere we have reviewed the theoretical implications of this relationship and shall not pursue it further here. But we believe that the dynamic time sense is intimately involved with the rise of Western civilization and the emergence of an industrial and capitalistic social order.

REFERENCES

1. KNAPP, R. H., & GARBUTT, J. T. Time imagery and the achievement motive. *J. Personal.*, 1958, **26**, 426-434.
2. MCCLELLAND, D. C., ATKINSON, J. W., CLARKE, R. A., & LOWELL, E. The Achievement Motive. New York: Appleton, 1953.
3. OSGOOD, C. E. The nature and measurement of meaning. *Psychol. Bull.*, 1952, **49**, 197-237.

Psychological Laboratory
Wesleyan University
Middletown, Connecticut

A STUDY OF PREJUDICE IN NEGRO AND WHITE COLLEGE STUDENTS*¹

*Department of Psychology, Emory University*²

LUIS PROENZA AND BONNIE R. STRICKLAND

A. INTRODUCTION

Despite radical social upheaval in regard to Negro civil rights demands, there has been little contemporary research on prejudice and attitude change accompanying increased Negro-White activity. The paucity of research is such that it is still important to describe and identify prejudicial attitudes. The present study is an attempt to document intensity of prejudice among Southern and Northern college students, as well as to investigate significant methods of assessing prejudice.

In the main, previous research has supported the belief that Southern whites are considerably more prejudiced toward the Negro than are Northern whites (7, 9, 10, 12, 13). Moreover, through continuous research since 1925, Bogardus has shown that whites of all regions have rated Negroes consistently low on the social-distance scale; i.e., white persons have felt a considerable amount of social distance between themselves and the Negro. However, some change is taking place, as is evidenced by data (obtained by Bogardus in 1956) that showed whites to rate Negroes as considerably less socially distant than did ratings 10 years earlier.

The ratings made by Negroes generally show much less social distance and prejudice toward whites than do ratings by whites on their attitudes toward Negroes (3, 14). However, in data reported in 1953 by Gray and Thompson (6) and in 1963 by Fagan and O'Neil (5), the ratings of Negro students toward a combination of different ethnic groups show more social distance than do ratings of the same groups by whites. In other words, Negroes show more social distance in regard to different ethnic groups than do whites; but they generally show less prejudice toward whites than whites do toward Negroes.

In regard to sex differences, Pettigrew (9) reported that women in the

* Received in the Editorial Office, Provincetown, Massachusetts, on May 21, 1964. Copyright, 1965, by The Journal Press.

¹ Portions of this study were presented at the Southeastern Psychological Association meetings, Gatlinburg, Tennessee, April, 1964.

² This research was supported by funds available from the Emory University Research Committee.

South are significantly more prejudiced than are males in the South; while no significant difference was found between sexes in the North. Bogardus (4) also reported that women show more social distance toward the Negro than do males; however, in an earlier study, Sims and Patrick (13) reported that women were significantly more favorable toward the Negro than were males.

Rating scales of prejudice in the past, for the most part, have made use of questionnaires or interviews that require an overt response in regard to some interaction with the particular ethnic group rated. Another approach to the study of prejudice may be that of Osgood and his colleagues (8), who have experimented with the assessment or measurement of meaning of certain concepts. Osgood has suggested that an experimenter may measure both the direction and the intensity of attitudes by using a set of scales that have high loadings on an evaluative factor across concepts. In respect to the use of the semantic differential to describe prejudice, Osgood *et al.* (8) have reported a comparison of the semantic-differential and Thurstone scales designed to measure the concepts "The Negro," "The Church," and "capital punishment." Reliability coefficients were shown to be high and equivalent; and, in each case, a significant correlation between the semantic-differential scale and the Thurstone scale was found. A study by Tannenbaum—reported by Osgood *et al.* (8)—also showed significant correlations between Bogardus social-distance ratings and semantic-differential ratings of Germans, Chinese, and Hindus.

To study regional, race, and sex influences on prejudice, the authors in the present study used both a modified form of the Bogardus social-distance scale and semantic-differential ratings for the concepts Negro, white, integration, and segregation.

B. METHODOLOGY

Subjects were members of introductory-psychology classes. Instructors agreed to have their classes participate in psychological experimentation, but students had no knowledge of the study until they were asked in class to complete the rating scales. All subjects cooperated.

The population tested consisted of 35 white males and 24 white females at a large university located in Ohio; 36 white males and 33 white females at a relatively small private university in Atlanta, Georgia; 23 white males and 27 white females at a midtown state college in Atlanta, Georgia; and 61 Negro males and 18 Negro females at a small private college in Atlanta, Georgia. The Southern schools were not integrated when the samples were collected. In each case the scales were administered by the instructors of the respective classes. Subjects were not asked to sign their names on the question-

naires, but were required to give age and sex. Final groups consisted of 79 Negro students, 119 white students in Southern schools, and 59 white students in a Northern school.

A booklet was constructed of a modified Bogardus social-distance scale and semantic-differential scales for the concepts Negro, white, integration, and segregation. Semantic-differential scales were the same for all subjects. For white students, the one form of the Bogardus scale used consisted of seven statements ranging from "I would have Negroes live in my country" to "I would marry a Negro." The Negro students rated whites on the same statements. Instructions for the Bogardus scale were as follows:

For each of the seven statements below, indicate either True or False by checking the appropriate blank. Your answer should be in relation to the white [Negro] population in general, and not in relation to any one isolated individual you have known. Try to answer as quickly as possible, since first impressions aid in the degree of validity of the question.

Instructions for the semantic differential were:

In each of the next four pages, you will find a set of scales which are constructed to measure a particular concept at the top of each page. Rate each concept against each of the adjectives according to what the concept means to you.

If you feel the concept you are rating is very personally related to either of the adjectives, then place an *X* in the appropriate direction.

If the adjectives should be equally related or completely irrelevant, then place an *X* in the central or neutral space.

Relatedness should be considered as follows.—*very: quite: slightly: neutral: slightly: quite: very:—*

Please work with speed and check every item.

Be sure you understand the instructions above.

The semantic differential was composed of 20 pairs of bipolar adjectives high on the evaluative factor according to Osgood (8). Such adjectives as good—bad and perfect—imperfect were used.

C. RESULTS

No significant differences in means and standard deviations were found between ratings of students of the two white Southern schools. Consequently, all white Southern student's ratings were combined for analysis.

Means and standard deviations of the Bogardus scale and semantic-differential ratings by Negro students, white students in the Northern schools, and white students in Southern schools (males and females) are shown in Table 1.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF THE BOGARDUS SCALE
AND THE SEMANTIC-DIFFERENTIAL RATINGS

Subjects, sex, and parameter	N	Bogardus scale	Semantic-differential scale			
			White	Negro	Segre- gation	Inte- gration
Negro						
Male	61					
Mean		6.46	3.50	3.37	5.27	2.52
SD		.67	.67	.67	1.05	.65
Female	18					
Mean		6.28	2.95	3.02	5.33	2.56
SD		.56	.48	.69	.71	.66
Southern white						
Male	59					
Mean		4.75	3.42	4.18	4.33	3.79
SD		1.83	.52	.52	.96	.88
Female	60					
Mean		4.83	3.45	3.88	4.55	3.45
SD		1.08	.42	.39	.90	.78
Northern white						
Male	35					
Mean		5.11	3.25	3.95	4.62	3.54
SD		1.33	.50	.66	.67	.76
Female	24					
Mean		5.83	3.38	3.74	4.53	3.54
SD		.92	.60	.65	1.03	.82

A two-by-three analysis of variance of the Bogardus ratings gives a highly significant F between schools [see Table 2]. Computation of t tests shows that the Negroes exhibit significantly less social distance toward whites than Southern whites exhibit toward Negroes. No significant differences were found between Northern and Southern male ratings on the Bogardus social-distance scale; however, white females in the Northern school show significantly less social distance toward Negroes than do white females in the Southern schools ($t = 4.15$, $p < .001$).

On the semantic-differential ratings of the four concepts Negro, white, integration, and segregation, t tests reveal no significant differences between

TABLE 2
ANALYSIS OF VARIANCE FOR THE BOGARDUS SCALE

Source of variation	df	Mean square	F
Sex	1	2.03	1.64
Race	2	62.95	50.77*
Interaction			
Sex-by-race	2	3.02	2.44
Within	251	1.24	

* $p < .01$.

white students of the Northern school and white students of the Southern schools; consequently, the semantic differential data of white students were combined for analysis. A two-by-two analysis of variance, race by sex, was computed for each of the concepts [see Tables 3 and 4]. Significant differences

TABLE 3
ANALYSIS OF VARIANCE FOR THE SEMANTIC-DIFFERENTIAL
CONCEPTS "NEGRO" AND "WHITE"

Source of variation	df	Concept			
		Negro		White	
		Mean square	F	Mean square	F
Sex	1	.79		.33	1.04
Race	1	25.61	68.84*	.03	
Interaction					
Sex-by-race	1	3.78	10.16*	4.03	12.67*
Within	253	.37		.32	

* $p < .01$.

TABLE 4
ANALYSIS OF VARIANCE FOR THE SEMANTIC-DIFFERENTIAL
CONCEPTS "SEGREGATION" AND "INTEGRATION"

Source of variation	df	Concept			
		Segregation		Integration	
		Mean square	F	Mean square	F
Sex	1	.39		.40	
Race	1	34.31	39.39*	62.36	103.07*
Interaction					
Sex-by-race	1	.20		1.83	
Within	253	.87		.60	

* $p < .01$.

between races were found for the ratings of Negro, segregation, and integration—with Negro students being more favorable toward the concepts Negro and integration and less favorable toward segregation than were whites. Negro students are more favorable toward the concept "White" than are white students toward the concept Negro. On all four concepts, no significant differences were found between sexes; however, significant interactions of sex by race were found for the concepts Negro and White.

To ascertain the relationship between the Bogardus social-distance scale and the semantic-differential ratings, Pearsonian correlations were computed [see Table 5]. In general, the Bogardus scale relates positively with semantic ratings of segregation and negatively with ratings of integration: i.e., high social distance, favorable evaluation of segregation and less favorable evalu-

TABLE 5
INTERCORRELATIONS OF BOGARDUS SCALE AND SEMANTIC-DIFFERENTIAL RATINGS

Subjects and sex	N	Semantic-differential concept			
		White	Negro	Segregation	Integration
Negro					
Male	61	.38**	.17	.29*	— .32*
Female	18	.30	.31	.23*	— .06
Southern white					
Male	59	.32*	— .23	.64**	— .48**
Female	60	.06	— .22	.51**	— .42**
Northern white					
Male	35	.09	— .10	.37*	— .17
Female	24	.51**	— .52*	.79**	— .58**

* $p < .05$.

** $p < .01$.

ation of integration. No consistent, significant correlations were found between the Bogardus scale and semantic ratings of the concepts Negro and White.

Table 6 gives the intercorrelations of semantic-differential scales and shows significant relationships between the concepts Negro and White and between the concepts integration and segregation (the only nonsignificant r is for Negro females, a group of 18 subjects). Ratings of Negro students for the

TABLE 6
INTERCORRELATIONS OF RATINGS ON SEMANTIC-DIFFERENTIAL SCALE

Concept and subject group	N	White	Concept Segregation	Integration
Negro (concept)				
White (subjects)				
Male	94	.25**	— .48**	.55**
Female	84	.38**	— .06	.33**
Negro (subjects)				
Male	61	.43**	— .04	.29*
Female	18	.49*	— .34	— .11
White (concept)				
White (subjects)				
Male	94		.05	.18*
Female	84		.17	.04
Negro (subjects)				
Male	61		.21*	.22*
Female	18		.00	.42**
Segregation (concept)				
White (subjects)				
Male	94			— .59**
Female	84			— .71**
Negro (subjects)				
Male	61			— .44**
Female	18			.00

* $p < .05$.

** $p < .01$.

concept White and white students for the concept Negro related significantly and positively with the ratings of the concept integration. While correlations are significant, it should be noted that they are relatively low.

D. DISCUSSION

In general, the Bogardus and semantic-differential findings suggest that Negro-white prejudicial attitudes contain differing elements of ethnic acceptance. The expected differences between Northern and Southern college student ratings were significant only in the case of the Bogardus social-distance scale. That significance appears to be due primarily to the finding that females in the Northern school showed less social distance toward the Negro than females in the Southern school. The semantic-differential ratings of the white students show no significant differences in regard to either region or sex. The Negro students, however, were considerably more emphatic in their favorable evaluation of integration and disfavor of segregation than were the white students.

In the main, the use of evaluative-differential scales appears to give a more definitive picture of underlying attitudes than do relatively simple measures of prejudice. The use of the semantic differential seems to be appropriate for a study of prejudice because of the consistent, significant relationships between certain of the concepts and ratings on the Bogardus scale. The present study used semantic-differential scales composed primarily of adjectives high on the evaluative factor. Further research may be pointed toward the use of total semantic space and the relationship of activity, potency, and evaluative factors as discussed by Osgood *et al.* (8).

The present study has documented meaningful differences between attitudes of white and Negro students in regard to prejudice; however, one should note certain limitations that pose problems in regard to the generalization of the results. First, the number of students tested was small, especially the group of Negro females ($N = 18$). Second, the schools might be quite different in regard to the kinds of students that would be in attendance. The Northern school was an integrated, large, state-supported University. The Southern white schools were a state-supported college and a private church-related University. The Negro colleges were private, church-related schools. Finally, the generalization of results to persons other than college students might be inappropriate. The attitudes of college students, both Negro and white, in a large Southern urban area, which has witnessed relatively progressive increased Negro-white activity can be expected to be different from the attitudes of Southern adults or the attitudes of Southern college students in

rural areas or other communities. It will be important to assess other samples as well as to identify changes in prejudice, if any occur, with increased civil-rights activities and desegregation of schools. It would also be interesting to assess Negro students who are attending Northern schools in order to compare their evaluations with those of Negro students in the deep South.

E. SUMMARY

A modified Bogardus social-distance scale and semantic-differential scales for the concepts Negro, White, integration and segregation were administered to 79 Negro students and 119 white students attending colleges in the South and 59 white students attending a large University in the North. Differences in ratings made by the white students of the Northern and Southern schools were significant only in the case of the Bogardus scale. Significant differences between races were found for the semantic-differential ratings of Negro, segregation and integration—with Negro students being more favorable toward the concepts Negro and integration and less favorable toward segregation than were whites. Negro students were more favorable toward the concept White than were white students toward the concept Negro. Significant correlations were found between the Bogardus social-distance and semantic-differential ratings of integration and segregation but not with ratings of Negro and White.

REFERENCES

1. BOGARDUS, E. S. Measuring social distances. *J. Appl. Sociol.*, 1925, **9**, 299-308.
2. ———. Changes in racial distances. *Internat. J. Opin. & Attitude Res.*, 1947, **1**, 55-62.
3. ———. Racial distance changes in the United States during the past thirty years. *Sociol. & Soc. Res.*, 1958, **43**, 127-135.
4. ———. Race reaction by sexes. *Sociol. & Soc. Res.*, 1959, **43**, 439-441.
5. FAGEN, J., & O'NEIL, M. A comparison of social distance scores among college student samples. *J. Soc. Psychol.* (in press).
6. GRAY, J. S., & THOMPSON, A. H. The ethnic prejudices of white and Negro college students. *J. Abn. & Soc. Psychol.*, 1953, **48**, 311-313.
7. HYMAN, H. H., & SHEATSLEY, P. B. Attitudes toward desegregation. *Scientific Amer.*, 1956, **195**, 35-39.
8. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The Measurement of Meaning*. Urbana: Univ. Illinois Press, 1957.
9. PETTIGREW, T. F. Regional differences in anti-Negro prejudice. *J. Abn. & Soc. Psychol.*, 1959, **59**, 28-36.
10. PROTHRO, E. T. Ethnocentrism and anti-Negro attitudes in the deep South. *J. Abn. & Soc. Psychol.*, 1952, **47**, 105-108.
11. PROTHRO, E. T., & MILES, O. K. A comparison of ethnic attitudes of college students and middle class adults from the same state. *J. Soc. Psychol.*, 1952, **36**, 53-58.

12. SAMELSON, B. The patterning of attitudes and beliefs regarding the American Negro: An analysis of public opinion. Unpublished Doctoral dissertation, Radcliffe College, Cambridge, Mass., 1945.
13. SIMS, V. M., & PATRICK, J. R. Attitude towards the Negro of Northern and Southern college students. *J. Soc. Psychol.*, 1936, **7**, 192-204.
14. WEBSTER, S. W. The influence of interracial contact on social acceptance in a newly integrated school. *J. Educ. Psychol.*, 1961, **52**, 292-296.

Department of Psychology

Emory University

Atlanta, Georgia 30322

STUDIES IN SOCIAL ATTITUDES: I. VARIATIONS IN
LATITUDES OF ACCEPTANCE AND REJECTION
AS A FUNCTION OF VARYING POSITIONS
ON A CONTROVERSIAL SOCIAL ISSUE*

Department of Psychology, American University of Beirut

LUTFY N. DIAB¹

A. INTRODUCTION

The terms *latitude of acceptance* and *latitude of rejection* have been recently introduced in the social-psychological literature mainly as a result of Sherif's and Hovland's researches and publications (5, 6, 7, 8). "Latitude of acceptance" is defined as consisting of "those positions on an issue that the individual finds most acceptable, plus other positions also acceptable to him" (5, p. 204); while "latitude of rejection" consists of "those positions on the same issue that he finds most objectionable, plus other objectionable positions on the issue" (5, pps. 204-205). Sherif proposes that the use of the foregoing concepts represents the individual's stand on an issue much more meaningfully than the usual single score or average yielded by most of the conventional attitude scales (5, 7). Consequently, a number of studies have been undertaken (e.g., 1, 2, 4) for the assessment of latitudes of acceptance and rejection, pointing the usefulness of such a conceptualization of social attitudes in studying effects of communication and attitude change. The main purpose of the present study is to retest the following two hypotheses:

Hypothesis 1. The latitude of rejection of individuals committed to an extreme stand on an ego-involving issue is greater than the latitude of rejection of individuals with a moderate position on the same issue.

Hypothesis 2. The latitude of rejection of individuals committed to an extreme stand on an ego-involving issue will be relatively greater than their latitude of acceptance (5, p. 205).

In addition to retesting the foregoing hypotheses with a different controversial social issue (namely, Arab unity) in a different culture, the present study investigates two additional points:

* Received in the Editorial Office, Provincetown, Massachusetts, on May 25, 1964. Copyright, 1965, by The Journal Press.

¹ The publication of this study was made possible by a grant to the School of Arts and Sciences at the American University of Beirut by the Rockefeller Foundation. Dr. Ernest Dalrymple-Alford's statistical advice and comments and Miss Suad Kaylani's research assistance are both greatly appreciated by the author.

1. Further refinements in the measurement technique for the assessment of latitudes of acceptance and rejection, as suggested by Sherif and Hovland (1, 7) and as will be explained in the section on procedure.

2. The study of latitudes of acceptance and rejection of individuals on more than one dimension of the same issue.

The foregoing points will enable us to test our third hypothesis, which may be stated as follows:

Hypothesis 3. Individuals holding an extreme stand on an ego-involving issue show greater consistency among various dimensions of their attitude toward the issue than do individuals holding moderate positions on the same issue.

B. METHOD

1. Subjects

The sample used in this study consisted of 260 undergraduate Arab students at the American University of Beirut. The median and mean ages for the total sample were 20 and 20½ years respectively. Of this sample, 198 were males and 62 were females. In terms of religious membership there were 136 Moslems, 111 Christians, and the remaining 13 subjects either indicated no religious affiliation or else belonged to a religious sect known as Druze. As to nationality, 124 were Lebanese and 94 were Jordanians or Palestinians; the remaining 42 belonged to other Arab countries: mainly, Syria, Sudan and Iraq.

2. Procedure

Data pertinent to the hypotheses and points were obtained from the subjects' responses to a questionnaire on "Arab unity," a highly ego-involving and controversial social issue in the Arab World.² The administration of the questionnaire took place in regular classroom sessions and lasted approximately 15 minutes. Other than securing background information on the subjects, the questionnaire consisted of two sets of nine statements each. One set was relevant to the *good-bad* dimension of Arab unity; the other represented the *possible-impossible* dimension of the same issue. The nine statements for each set were chosen after preliminary tests on 30 individuals known to the experimenter to represent different stands on this issue. Thus one set of nine statements represented a continuum of clearly differentiated positions on one dimension (the *evaluative* dimension); while the other set of nine statements represented a continuum of different positions on the second dimension (the

² Data for this study were collected during the period extending through the months of January, February, and March, 1963.

possibility dimension). For each dimension, the complete set of nine statements was mimeographed on each of four sheets of paper.

On the first sheet, subjects indicated the statement "most acceptable" to them; on the second, they indicated other statements also acceptable to them. The positions checked on these two sheets provided the data on their latitudes of acceptance. On the third sheet, subjects checked the statement "most objectionable" to them, and on the last sheet they checked other objectionable statements. The checkings on these last two pages yielded data on their latitudes of rejection (5, p. 206).

The above procedure and instructions were repeated for both sets of nine statements. The middle and extreme statements on the "evaluative" dimension are reproduced below:

A. The achievement of Arab unity is absolutely essential from all angles in the Arab people's interests.

E. From the point of view of the Arab people's interests, it is hard to decide whether Arab unity would be advantageous or disadvantageous.

I. The achievement of Arab unity is completely detrimental from all angles in the Arab people's interests.

Similar statements on the "possibility" dimension are also reproduced below:

A. There is no doubt whatsoever that Arab unity will definitely be achieved in the future.

E. It is hard to tell whether Arab unity will or will not be achieved in the future.

I. There is no doubt whatsoever that Arab unity will never be achieved at any time in the future.

Though the procedure for testing the two main hypotheses of this study is essentially the same as that followed by Sherif and Hovland (5, 7), it differs from theirs in that the hypotheses were tested on more than one dimension of the same issue (the "evaluative" and the "possibility" dimensions of Arab unity) and in trying out a refinement to the Sherif-Hovland measurement technique. This refinement in measurement consisted of asking the subjects to indicate how strongly they felt about every statement they checked in any way (i.e., "most acceptable," "also acceptable," "most objectionable," and "also objectionable" statements), by placing before the statement one of the following: VS (Very strongly), S (Strongly), or M (Mildly). Such a modification of the Sherif-Hovland technique was intended (a) to help reveal individuals strongly committed to a moderate or neutral stand and (b) to differentiate further between individuals with apparently similar extreme stands. Such fine differentiations among various positions on the issue could not be made with the original Sherif-Hovland technique.

C. RESULTS

The main results are summarized in Tables 1 through 5. Subjects were classified into extreme pros (those selecting statements A, B, or C as "most acceptable" to them), moderates (those selecting statements D, E, or F as "most acceptable" to them), and extreme anti-Arab unity subjects (those selecting statements G, H, or I as "most acceptable" to them). As shown in Table 1, all three subsamples of subjects reject significantly more items or

TABLE 1
DIFFERENCES BETWEEN MEAN NUMBER OF "OBJECTIONABLE" AND "ACCEPTABLE"
STATEMENTS TO SUBJECTS WITH EXTREME AND MODERATE
POSITIONS ON THE ARAB-UNITY ISSUE

Subject's position	N	a Mean number of items objectionable	b Mean number of items acceptable	a-b Difference	t
Evaluative dimension					
Extreme pro	179	3.46	2.70	0.76	7.465**
Moderate	53	2.60	2.26	0.34	2.897**
Extreme anti	24	3.21	2.54	0.67	2.320*
Possibility dimension					
Extreme pro	193	3.73	2.31	1.42	11.455**
Moderate	43	2.41	2.46	-0.05	0.298
Extreme anti	20	2.75	2.55	0.25	0.689

* Significant at the .05 level.

** Significant at the .01 level.

statements than they accept on the evaluative dimension. Further, there are no significant differences between the mean number of statements rejected by these subsamples on the evaluative dimension. These findings, with the exception of those pertaining to moderate subjects, are consistent with the first two hypotheses. As to the possibility dimension, even though all of the differences in means were in the expected directions, only the extreme pro-Arab unity subjects rejected significantly more items or statements than they accepted [see Table 1] and, in addition, they rejected a significantly greater number of statements than did the moderate ($t = 4.98$, $p < .01$) and the extreme anti-Arab unity subjects ($t = 2.15$, $p < .05$), with these last two groups not differing significantly in this respect. These findings on the possibility dimension, with the exception of those pertaining to the extreme anti-Arab unity subjects, were in line with expectations.

To get further clarification of the above results concerning an individual's stand and variations in his latitudes of acceptance and rejection, the data were regrouped as shown in Tables 2 and 3 by classifying the subjects not only in

TABLE 2
DIFFERENCES BETWEEN MEAN NUMBER OF "OBJECTIONABLE" AND
"ACCEPTABLE" STATEMENTS FOR EXTREME AND MODERATE
SUBJECTS ON THE ARAB-UNITY ISSUE

Subject's position and degree of acceptance	N	a Mean number of items objection- able	b Mean number of items acceptable	a-b Differ- ence	t
Evaluative dimension					
Extreme pro-Arab unity					
Very strong and strong	174	3.71	2.72	0.99	7.417*
Mild	5	3.00	2.20	0.80	1.000
Moderate					
Very strong and strong	44	2.75	2.27	0.48	3.878*
Mild	9	1.88	2.22	-0.34	1.417
Possibility dimension					
Extreme pro-Arab unity					
Very strong and strong	185	3.79	2.32	1.47	11.579*
Mild	8	2.25	2.00	0.25	0.798
Moderate					
Very strong and strong	31	2.45	2.54	-0.09	0.485
Mild	12	2.33	2.25	0.08	0.364

* Significant at the .01 level.

TABLE 3
DIFFERENCE BETWEEN MEAN NUMBER OF "OBJECTIONABLE" STATEMENTS BY
EXTREME AND MODERATE SUBJECTS ON THE ARAB-UNITY ISSUE

Subject's position on Arab unity	Evaluative dimension		Possibility dimension	
Extreme pro (very strong and strong)	174	3.71	185	3.79
Extreme pro (mild)	5	3.00	8	2.25
Difference		0.71		1.54
t		1.09		2.47*
Moderate (very strong and strong)	44	2.75	31	2.45
Moderate (mild)	9	1.88	12	2.33
Difference		0.67		0.12
t		2.70*		0.31

* $p < .01$.

terms of the statement or position "most acceptable" to them but also in terms of the degree to which they accept their "most acceptable" position (i.e., very strong, strong, or mild acceptance of the "most acceptable" position). It will be seen from Table 2 that the extreme pro-Arab unity subjects who expressed "mild" acceptance of their stand did not reject significantly more items than they accept, a phenomenon that holds true for both dimensions. Also, while the moderates who expressed "mild" acceptance of their

stand on the evaluative dimension accepted more items than they rejected, the moderates who expressed "very strong" or "strong" acceptance of their moderate stand rejected significantly more items than they accepted. This latter finding concerning the moderate subjects on the evaluative dimension, though not in line with expectations, becomes understandable in the light of strength of commitment to the moderate stand. On the other hand, the moderates on the possibility dimension, as expected, did not show significant differences between their latitudes of acceptance and rejection, irrespective of the degree with which they upheld their stand [see Table 2]. Moreover, as shown in Table 3, it is clear that on the evaluative dimension the moderate subjects who felt "very strongly" or "strongly" about their position on the issue rejected significantly more items than the moderates who expressed "mild" acceptance of their position, a finding which does not apply to the moderates on the possibility dimension (though the differences were in the expected direction). Also, the extreme pro-Arab unity subjects who felt "very strongly" or "strongly" about their position on the possibility dimension rejected significantly more items than the extreme pro subjects who expressed "mild" acceptance of their position, a finding which does not apply to the extreme pro-Arab unity subjects on the evaluative dimension, even though the differences were in the expected directions [see Table 3]. Thus, it may be stated here that the degree of commitment to a stand may be in itself an important factor affecting latitudes of acceptance and rejection.

To test the third hypothesis of this study (namely, that subjects with extreme stands on an issue show greater consistency among various dimensions of their attitude than do moderate subjects) Table 4 was constructed. This Table shows that, of the total number of extreme *pro* subjects on the evaluative dimension, about 94 per cent are also extreme *pro* on the possibility dimension; however about one-half of the moderates and half of the extreme anti-Arab unity subjects on the evaluative dimension are also, respectively, moderates and extreme anti-Arab unity on the possibility dimension. A chi-

TABLE 4
NUMBER AND PER CENT OF SUBJECTS CHECKING VARIOUS POSITIONS
AS "MOST ACCEPTABLE" ON THE EVALUATIVE DIMENSION
OF THE ARAB-UNITY ISSUE

Possibility dimension	Extreme pro		Evaluative dimension		Extreme anti	
	N	%	N	%	N	%
Extreme pro	168	93.86	22	41.50	3	12.50
Moderate	9	5.03	25	47.18	9	37.50
Extreme anti	2	1.11	6	11.32	12	50.00
Total	179	100.00	53	100.00	24	100.00

square test based on the frequencies in Table 4 yields a value of 144.530, significant beyond the .001 level (for four degrees of freedom) and implies that the evaluative and possibility dimensions relevant to the Arab unity issue are not independent of each other. In other words, the stand of a subject on one dimension is likely to be the same on the other dimension. This is particularly the case for the extreme *pro* subjects, less so for the extreme anti-Arab unity subjects, and least for the moderates. These findings give general support to the third hypothesis of this study, though it is clear that extreme anti-Arab unity subjects do not differ significantly from moderate subjects. It seems that the greatest consistency between both dimensions holds true mainly for the extreme pro-Arab unity subjects.

As stated earlier in the procedure section, subjects were asked to indicate how strongly they feel about the various statements they accept or reject. Tabulation of the subjects' responses in terms of various degrees of acceptance for the "most acceptable" position is presented in Table 5. For the evaluative dimension, about 97 per cent of the extreme *pro* and all of the extreme anti-Arab unity subjects stated that they felt either "very strongly" or "strongly" about the positions they checked as "most acceptable," which are Positions A, B, and C and G, H, and I respectively. However, the moderate subjects also expressed similar feelings on the evaluative dimension, for about 83 per cent of them stated that they felt either "very strongly" or "strongly" about the positions they chose as "most acceptable" to them (i.e., Positions D, E, or F), and only about 17 per cent of them felt "mildly" about their "most acceptable" position. A chi-square test based on the frequencies shown in Table 5 for the evaluative dimension yields a value of 39.737, which is significant beyond the .001 level (for four degrees of freedom). Partitioning this chi-square value (3) shows that the extreme pro-Arab unity subjects, in comparison to the moderates, check either "very strongly" or "strongly" significantly more than they check "mildly" to represent the strength of their feeling about their "most acceptable" position ($\chi^2 = 15.920$, significant beyond the .001 level for one degree of freedom). Also, the extreme anti-Arab unity subjects, in contrast to moderate subjects, check either "very strongly" or "strongly" significantly more than they check "mildly" to indicate the strength of their feeling about their "most acceptable" position ($\chi^2 = 9.200$, significant beyond the .01 level for one degree of freedom). However, as expected, the extreme *pro* and the extreme anti-Arab unity subjects do not differ significantly in the degree with which they accept their "most acceptable" positions ($\chi^2 = 0.318$, not significant for one degree of freedom).

As to the possibility dimension, it is clear from Table 5 that about 96 per

TABLE 5
NUMBER AND PER CENT OF SUBJECTS CHECKING VARIOUS POSITIONS AS "MOST ACCEPTABLE" ON THE ARAB-UNITY ISSUE
SHOWING VARIOUS DEGREES OF ACCEPTABILITY

Degree of acceptability	Evaluative dimension						Possibility dimension					
	Extreme pro <i>N</i>	%	Moderate <i>N</i>	%	Extreme anti <i>N</i>	%	Extreme pro <i>N</i>	%	Moderate <i>N</i>	%	Extreme anti <i>N</i>	%
Very strongly	137	76.54	18	33.96	17	70.83	122	63.22	18	41.87	14	70.00
Strongly	37	20.67	26	49.05	7	29.17	63	32.64	13	30.23	5	25.00
Mildly	5	2.79	9	16.99	0	0.00	8	4.14	12	27.90	1	5.00
Total	179	100.00	53	100.00	24	100.00	193	100.00	43	100.00	20	100.00

cent of the extreme *pro* and 95 per cent of the extreme anti-Arab unity subjects stated that they felt either "very strongly" or "strongly" about the positions they checked as "most acceptable" to them. However, only about 72 per cent of the moderate subjects felt the same way about their own "most acceptable" position, and the remaining 28 per cent of moderates stated that they felt "mildly." A chi-square test based on the frequencies presented in Table 5 for the possibility dimension yields a value of 27.855, which is significant beyond the .001 level (for four degrees of freedom). Partitioning this chi-square value (3) shows that the extreme pro-Arab unity subjects, in contrast to the moderates, check either "very strongly" or "strongly" significantly more than they check "mildly" to indicate the strength of their feeling about their "most acceptable" position ($\chi^2 = 26.340$, significant beyond the .001 level for one degree of freedom). Also, the extreme anti-Arab unity subjects, in comparison to the moderates, check "very strongly" or "strongly" significantly more than they check "mildly" to indicate the strength of their feeling about their stand on the issue ($\chi^2 = 9.509$, significant beyond the .01 level for one degree of freedom). However, as expected, the extreme *pro* and the extreme anti-Arab unity subjects on the possibility dimension do not differ significantly in the degree with which they accept their stand ($\chi^2 = 0.175$, not significant for one degree of freedom).

D. DISCUSSION

The results of this study may best be understood by comparing them with previous findings on which they were based, mainly represented by Sherif's and Hovland's work (1, 5, 6, 7). In line with the findings of these two authors, it was found in this study that individuals upholding an extreme favorable stand on an issue have latitudes of rejection greater than the latitude of rejection of subjects upholding moderate positions on the same issue, though the difference is significant only on the possibility dimension. Thus, it is clear that these results pertaining to subjects with an extremely favorable stand support the first hypothesis of this study and confirm previous findings by Hovland and Sherif (1, 7). On the other hand, the latitude of rejection of the extremely anti-Arab unity subjects is also greater than that of the moderates; though the difference is not significant on either of the two dimensions. Thus, the results pertaining to subjects with an extremely unfavorable stand also support, though not significantly, the first hypothesis of this study and are in line with previous findings.

However, in line with the first hypothesis, the extremely *pro* and the extremely anti-Arab unity subjects were not expected to differ significantly

from each other in their latitudes of rejection. Our results show that, on the possibility dimension, the mean number of items objectionable to the extreme *pro* subjects is significantly greater than the mean number of items objectionable to the extreme anti-Arab unity subjects. Though this finding is not in line with the implications of the first hypothesis, it may be better understood when considered in relation to the nature of the particular social issue under consideration: namely, the issue of Arab unity. It may be pointed out here that the operation of certain reality checks have caused the extremely anti-Arab unity subjects to be reluctant in rejecting the possibility of unity among the Arab countries. For one thing, Arab unity materialized in 1958 when the United Arab Republic (U.A.R.) was created. Though this unity did not last till the time of this study, the present data were collected at a time when in Cairo talks for Arab unity and the formation of a new U.A.R. were about to be held among Egypt, Iraq, and Syria. Thus, though the extremely anti-Arab unity subjects expressed through the positions they chose as "most acceptable" to them (i.e., Positions G, H, or I) that they felt Arab unity was impossible to achieve, yet they chose to remain noncommittal on a relatively large number of the remaining positions on the issue. It appears that subjects who take an extremely opposed stand to Arab unity seem to be more reluctant and unsure of the practicality of their position, when compared to subjects who show an extremely favorable stand. Or perhaps one may speculate in this connection about the possibility of different effects produced by positive and negative ego involvements and resulting also from the reality checks pertaining to the issue at hand. Or, finally, one may raise a question as to whether the most-extreme end positions on the issue of Arab unity (i.e., Positions A and I) are equally extreme. That is, could it be that Position I is not as extreme as Position A? Further research, utilizing larger samples of subjects, is needed in order to understand more fully the factors operating to produce these differences between the extreme *pro* and the extreme anti-Arab unity subjects.

Our results seem also to give clear support to our second hypothesis: namely, that individuals with extreme stands on an ego-involving issue have latitudes of rejection greater than their own latitudes of acceptance. Again, in line with Sherif's and Hovland's findings (1, 5, 7), it was found that individuals upholding an extreme favorable or unfavorable stand on the evaluative dimension of the Arab-unity issue tend to reject more positions than they accept [see Table 1]. However, on the possibility dimension of the same issue, this finding held true mainly for subjects with an extremely favorable stand (at Positions A, B, and C) and only partially for subjects with an

extremely unfavorable stand [see Table 1]. As shown in Table 1, subjects with extremely favorable as well as unfavorable stands reject significantly more items than they accept on the evaluative dimension. However, on the possibility dimension, though all the differences are in the expected directions, only those subjects with an extremely favorable stand rejected significantly more items than they accepted [see Table 1]. Here again is another instance of an unexpected difference between subjects with extremely favorable and those with extremely unfavorable stands toward Arab unity, resulting perhaps from different kinds of ego involvement or from the compelling demands of reality and suggesting the need for further research to clarify these discrepancies.

As a corollary of our second hypothesis, it was expected that subjects with moderate stands on the issue of Arab unity should not evidence significant differences between their latitudes of acceptance and their latitudes of rejection. As Table 1 shows, such an expectancy held true on the possibility dimension but not on the evaluative dimension. In fact, on the evaluative dimension, moderate subjects rejected significantly more items than they accepted [see Table 1], thus showing similar behavior to subjects with extreme stands. As stated earlier, though this finding is not in line with our expectations, it becomes understandable in the light of the subject's strength of commitment to the moderate stand. As Sherif and Hovland remark, "provision must also be made, however, for the occasional individual who strongly upholds a position in the middle segments and rejects positions at the extremes" (7, p. 190). The results of the breakdown of the sample in terms of the subjects' stand as well as the degree with which they are committed to that stand, presented in Tables 3 and 4, lead us to conclude that the degree with which a subject upholds his stand, regardless of whether it is a moderate or an extreme stand [see Tables 3 and 4], is in itself an important factor determining the relative sizes of his latitudes of acceptance and rejection.

Finally, our results tended to support our third hypothesis: namely, that subjects with extreme stands showed greater consistency among their attitude components than did moderate subjects [see Table 5]. However, the greatest consistency was exhibited by the extreme *pro* subjects, while the extreme anti-Arab unity and the moderate subjects did not differ appreciably from each other. Again, we evidence here the irregular behavior of those subjects that have an extremely unfavorable stand towards Arab unity, in contrast to the consistent behavior of the subjects who have an extremely favorable stand. It is obvious here that while the evaluative and the possibility dimensions of

Arab unity appear to be strongly related when subjects with an extremely favorable stand are considered, the relationship tends to weaken considerably in the case of moderate as well as extreme anti-Arab unity subjects. This relative lack of consistency in the case of subjects with extremely unfavorable stands, in comparison to that of subjects with extremely favorable stands, may be due to a variety of factors. For one thing, subjects with an extremely unfavorable stand towards Arab unity may either feel that Arab unity has some advantages though it appears impossible to achieve or they may feel that it is very bad but yet has some chances of being achieved. In other words, it seems that the relative lack of consistency in the case of the extreme anti-Arab unity subjects may be the result of their extreme stand on one dimension and, at the same time, their less extreme stand (i.e., upholding a moderate or even an extreme pro position) on the other dimension. In either case, it seems that such relatively inconsistent behavior is caused by the fact that extreme negative stand on both dimensions may not be very realistic. On the other hand, the extremely consistent behavior of individuals upholding an extremely favorable stand towards Arab unity may be, at the same time, due to blindness to and denial of reality on the part of these subjects, resulting from extreme positive ego involvement. As suggested earlier, there is the possibility of different effects produced by positive and negative ego involvements on an issue. However, the need for further research on these and on related points is indicated by our results.

E. SUMMARY

The purpose of the present study was to investigate variations in the latitudes of acceptance and rejection as a function of varying positions on a controversial social issue: namely, the issue of Arab unity.

The procedure used for testing the hypotheses of this study involved the administration of a questionnaire which was essentially similar to that used by Sherif and Hovland (5, 7). The sample consisted of 260 undergraduate Arab students at the American University of Beirut.

The study showed, in line with previous findings by Sherif and Hovland, (a) that the latitude of rejection of subjects upholding extreme stands is greater than the latitude of rejection of subjects upholding moderate stands and (b) that the latitude of rejection of subjects upholding extreme stands is greater than their own latitude of acceptance. However, the results pertaining to subjects upholding an extremely negative stand towards Arab unity were not fully in line with expectations. The explanation given to account for this discrepancy was based on reality demands relevant to the issue under

consideration. Furthermore, the results pertaining to subjects upholding a moderate stand also were not fully in line with our hypotheses. Such a discrepancy was mainly accounted for by taking the subject's strength of commitment to his moderate stand into consideration.

Finally, our results have shown that subjects with extreme stands exhibit greater consistency among their attitudinal dimensions than do moderate subjects. However, subjects upholding an extremely favorable stand towards Arab unity showed a much greater consistency than did subjects with an extremely unfavorable stand. This discrepancy was explained in terms of the possible differential effects of positive and negative ego involvements in relation to the social issue in question and to the operation of reality checks.

The need for further research, utilizing larger samples of subjects, is indicated. This is particularly essential in relation to findings on subjects upholding extremely negative stands toward Arab unity.

REFERENCES

1. HOVLAND, C. I., HARVEY, O. J., & SHERIF, M. Assimilation and contrast effects in reactions to communication and attitude change. *J. Abn. & Soc. Psychol.*, 1957, 55, 244-252.
2. LA FAVE, L., & SHERIF, M. Placement of items on a controversial social issue. In M. Sherif & C. I. Hovland, *Social Judgment*. New Haven: Yale Univ. Press, 1961.
3. MAXWELL, A. E. *Analysing Qualitative Data*. London: Methuen, 1961.
4. SHERIF, C. W. Social categorization as a function of latitude of acceptance and series range. *J. Abn. & Soc. Psychol.*, 1963, 67, 148-156.
5. SHERIF, M. Some needed concepts in the study of social attitudes. In J. G. Peatman & E. L. Hartley (Eds.), *Festschrift for Gardner Murphy*. New York: Harper, 1960.
6. ———. Conformity-deviation, norms, and group relations. In I. A. Berg & B. M. Bass (Eds.), *Conformity and Deviation*. New York: Harper, 1961.
7. SHERIF, M., & HOVLAND, C. I. *Social Judgment*. New Haven: Yale Univ. Press, 1961.
8. SHERIF, M., & SHERIF, C. W. *An Outline of Social Psychology*. New York: Harper, 1956.

Department of Psychology
American University of Beirut
Beirut, Lebanon

STUDIES IN SOCIAL ATTITUDES: II. SELECTIVITY IN
MASS COMMUNICATION MEDIA AS A FUNCTION
OF ATTITUDE—MEDIUM DISCREPANCY* 1

Department of Psychology, American University of Beirut

LUTFY N. DIAB

A. INTRODUCTION

A number of studies (1, 2, 4, 5) have shown that individuals upholding strong stands on an issue will have a "... *raised threshold of acceptance and lowered threshold of rejection*" (4, p. 202).

In line with the foregoing studies, the author speculated that an individual's stand on an issue should determine the kinds of mass communication media to which he exposes himself; hence this study is an attempt to provide an answer to the question that may be presented in the form of an hypothesis as follows: Individuals upholding an extreme positive or negative stand on an ego-involving issue are selective in the kinds of mass communication media to which they expose themselves, the selectivity being in the direction of media representing their own points of view to the exclusion of media expressing opposite points of view.

B. METHOD

1. *Subjects*

The same sample used in a recent study by the author (1) was utilized in this study. It consisted of 260 undergraduate Arab students at the American University of Beirut (198 males and 62 females) with a mean age of 20½ years. One hundred thirty-six were Moslems, and 111 were Christians; the remainder indicated no religious preference or belonged to a religious sect known as Druze. One hundred twenty-four were Lebanese and 94 were Jordanians or Palestinians; the remainder, mainly, were Syrians, Sudanese, and Iraqis.

2. *Procedure*

The questionnaire utilized in an earlier study by the author (1) contained questions relevant to the hypothesis of the present study. These questions asked the subject to indicate by name (a) the newspapers he reads or prefers to

* Received in the Editorial Office, Provincetown, Massachusetts, on May 25, 1964. Copyright, 1965, by The Journal Press.

¹ Thanks are extended to Miss Suad Kaylani for her assistance in tabulating the data.

read often and (b) the radio stations he listens to or prefers to listen to often. The questionnaire was entitled "A Questionnaire on Arab Unity" and was given in regular classroom sessions. It took approximately 15 minutes. As reported earlier (1), subjects were asked to express their stand on the good-bad (evaluative) dimension of Arab unity, by checking one out of nine statements (A to I) "most acceptable." Subjects checking A, B, or C positions (or statements) were classified as extreme pro-Arab unity; those checking D, E, or F were considered moderates; finally, subjects checking G, H, or I positions as "most acceptable" were classified as extreme anti-Arab unity.

Prior to collecting and tabulating the subjects' responses, a categorization of newspapers and radio stations available in the Arab Middle East area was prepared. This categorization was based on majority agreement among 10 judges, who were asked to place each of many mass media in one of the following categories: I. Pro-Arab unity newspapers and radio stations, II. "Mixed" newspapers and radio stations (i.e., neutral or foreign), III. Anti-Arab unity newspapers and radio stations. The criterion used for the placement of the media in any one of the categories was the placement of that medium in one category by six or more of the judges. The 10 judges (university professors, journalists, and graduate students) represented different positions on the issue of Arab unity.

C. RESULTS

The main results are presented in Tables 1 and 2. Table 1 shows that the results pertaining to the "reading" behavior of the extreme pro-Arab unity subjects and, particularly, the extreme anti-Arab unity subjects seem to give strong support to the hypothesis of this study. Though moderate subjects were not specifically referred to in our hypothesis, it was implied that they would show (in comparison to extreme subjects) less selectivity in the kinds of mass communication media to which they expose themselves; however, it appears

TABLE 1
NUMBER AND PER CENT OF SUBJECTS SHOWING PREFERENCES FOR
"NEWSPAPERS" EXPRESSING DIFFERENT POINTS OF VIEW
ON THE ARAB-UNITY ISSUE

Evaluative dimension	N	"Newspapers" preferred					
		Pro-Arab unity		Mixed		Anti-Arab unity	
		N	%	N	%	N	%
Extreme pro-Arab	166	57	34.34	78	46.99	31	18.67
Moderate	47	5	10.64	17	36.17	25	53.19
Extreme anti	24	1	4.17	3	12.50	20	83.33

TABLE 2
NUMBER AND PER CENT OF SUBJECTS SHOWING PREFERENCES FOR
"RADIO STATIONS" EXPRESSING DIFFERENT POINTS OF
VIEW ON THE ARAB-UNITY ISSUE

Evaluative dimension	N	"Radio stations" preferred					
		Pro-Arab unity		Mixed		Anti-Arab unity	
		N	%	N	%	N	%
Extreme pro-Arab	163	92	56.44	60	36.81	11	6.75
Moderate	49	10	20.41	29	59.18	10	20.41
Extreme anti	22	1	4.54	13	59.10	8	36.36

that the greatest proportion of moderate subjects (about 53 per cent) prefer to read anti-Arab unity papers, while only about 11 per cent read pro-Arab unity papers. A chi-square test, based on the frequencies shown in Table 1, yields a value of 54.712 (which is significant beyond the .001 level [$df = 4$]) and implies that the kinds of newspapers read by our subjects are not independent of the subjects' stands on the issue of Arab unity. Partitioning of the chi-square value (3) shows that the extreme pro-Arab unity subjects, in comparison to the extreme anti-Arab unity subjects, prefer to read pro-Arab unity papers significantly more than they read anti-Arab unity papers and vice versa ($\chi^2 = 30.450$, significant beyond the .001 level for one degree of freedom).

Table 2 shows that the findings relevant to the "listening" behavior of subjects upholding different stands on the Arab-unity issue give strong support to our hypothesis; thus, while the greatest proportion of the extreme pro-Arab unity subjects (about 56 per cent) prefer to listen often to pro-Arab unity radio stations, only about seven per cent listen to anti-Arab unity stations. Also, while only about five per cent of the extreme anti-Arab unity subjects listen to pro-Arab unity stations, about 36 per cent of them prefer to listen to anti-Arab unity stations. Finally, in line with the implications of our hypothesis, equal proportions of moderate subjects listen to pro-Arab unity as well as to anti-Arab unity stations. A chi-square test, based on the frequencies shown in Table 2, yields a value of 42.284, significant beyond the .001 level ($df = 4$) and implies that the kinds of radio stations preferred by our subjects are not independent of the subject's attitudes towards Arab unity. Partitioning this chi-square value (3) shows that the extreme pro-Arab unity subjects, in comparison to extreme anti-Arab unity subjects, prefer to listen to pro-Arab unity stations significantly more than they listen to anti-Arab unity stations and vice versa ($\chi^2 = 14.556$, significant beyond the .001 level for one degree of freedom).

D. DISCUSSION

The results of this study may be better understood if a point concerning procedure is clarified and made more explicit. It will be recalled that our subjects were asked to write the names of various newspapers and radio stations that they like or prefer. There is no sure way of telling the reasons that a subject may have for expressing preference to a particular newspaper or radio station, but a subject preferring to read a pro-Arab unity paper rather than an anti-Arab unity paper does so, most likely, because of the particular political orientation of the paper. However, we cannot be equally sure that a subject prefers to listen often to a particular radio station because of the latter's viewpoint on the issue of Arab unity. After all, there is a likelihood that the subject may prefer a certain radio station because it offers him the kind and variety of musical programs that he happens to like. Furthermore, there is no financial expenditure involved in listening to the radio, as in obtaining one's favorite newspaper (besides the fact that "reading" involves more active effort on the part of the individual than "listening"). In spite of the foregoing, one should not infer that the subjects' preferences for various media were not mostly determined by the political coloring of the media, particularly in the Arab Middle East. In addition, as mentioned earlier, the questionnaire was entitled "A Questionnaire on Arab Unity," which fact should put the subject in this particular context and have his preferences expressed mainly in relation to the issue in question.

In line with our hypothesis, subjects with an extreme positive stand towards Arab unity are highly selective in the kinds of radio stations they listen to (preferring pro-Arab unity and avoiding anti-Arab unity stations) and they also show considerable selectivity in the kinds of newspapers they read. However, about 66 per cent of the pro-Arab unity subjects [see Table 1] reported that they often read one or more "mixed" as well as anti-Arab unity papers. Yet, a further analysis of this result shows that the greatest proportion of these subjects who read "mixed" and anti-Arab unity papers (about 60 per cent of them) read only one particular newspaper (classified as anti-Arab unity by the majority of our 10 judges). This particular paper has special appeal to our pro-Arab unity subjects mainly as a result of a daily editorial which it includes, often revolving around the issue of Arab unity and the various problems that relate to it and which may be characterized as being analytical rather than critical of Arab unity. Thus, the apparent lack of strong selectivity in the kinds of newspapers read by our pro-Arab unity subjects is now more understandable in the light of this information. As our results show, about one-third of our pro-Arab unity subjects prefer to read pro-Arab unity

papers; while about two-thirds of them also reported preference for reading all other kinds of papers, including anti-Arab unity papers (of which one particular paper has special appeal because of its analytical daily editorial on the Arab World and its day-to-day problems). Thus, it appears that about two-thirds of our pro-Arab unity subjects are eager to learn also about various points of view other than their own. Being strongly and positively ego involved in the issue of Arab unity, the majority of our pro-Arab unity subjects are highly motivated to find out what problems stand in the way of achieving this unity. It is possible that such a finding would not be applicable to uneducated and nonuniversity pro-Arab unity subjects, where much greater selectivity in mass media would be expected. Such a point is worth investigating.

In line with our hypothesis, anti-Arab unity subjects show high selectivity in the kinds of newspapers they read, preferring anti-Arab unity newspapers and almost refusing to read pro-Arab unity papers. However, when it comes to radio station preferences, about 59 per cent of the anti-Arab unity subjects were found to prefer mostly "neutral" and "foreign" radio stations, with about 36 per cent reporting that they prefer to listen to anti-Arab unity stations. Thus, there is evidence for selectivity as well as withdrawal in the "reading" and "listening" behavior of our anti-Arab unity subjects. These findings about anti-Arab unity subjects may indicate a tendency on their part to run away from or avoid reality, caused by an underlying feeling of threat and insecurity that may accompany their minority-group status. In other words, we may have here an instance of perceptual defense contributing to the preservation and stability of one's already existent attitudes. On the other hand, it might be that the anti-Arab unity subjects, in contrast to the pro-Arab unity subjects, do not experience any great need for knowing and discovering more about the Arab unity issue. To them, Arab unity is impossible to achieve or disadvantageous to the Arab countries; hence there is no need to know anything more about it. Furthermore, one may add that most of the anti-Arab unity subjects seem to obtain their information from anti-Arab unity papers and from "neutral" and "foreign" as well as anti-Arab unity radio stations.

Finally, subjects with a moderate stand towards the Arab-unity issue were found to behave in a fashion highly similar to the anti-Arab unity subjects, particularly in connection with the kinds of newspapers they prefer to read. In fact, it was reported earlier that moderate subjects behaved very much like anti-Arab unity subjects in that they rejected significantly more items than they accepted and in that the great majority of them rejected the favor-

able end of the "evaluative" and "possibility" continua of Arab unity (1). Thus, the present finding on moderate subjects, though not fully in line with the implications of our hypothesis, confirms earlier findings (1) about the similarity in the reactions of moderate and anti-Arab unity subjects. Such a finding leads us to doubt the neutral position of so-called moderate subjects, as measured by direct questions.

E. SUMMARY

The purpose of the present study was to study selectivity in mass-communication media as a function of one's stand on a controversial social issue: namely, the issue of Arab unity.

The sample used in this study consisted of 260 undergraduate Arab students at the American University of Beirut. The subjects were given a questionnaire on Arab unity (1) on the basis of which their stand towards this issue was determined; also, subjects were asked to indicate their preferences for various newspapers and radio stations available in the Arab Middle East area.

In general, the results demonstrate that subjects with an extreme stand on the issue of Arab unity (particularly those with an extreme negative stand) show high selectivity in the kinds of mass communication media to which they expose themselves, the selectivity being in the form of preferences for media (i.e., newspapers and radio stations) expressing points of view similar to their own. However, subjects with a moderate stand on the issue of Arab unity also showed behavior similar to that of the extreme anti-Arab unity subjects. This result raises questions as to the adequacy of the attitude-measuring technique that was employed in this and in similar previous studies in the assessment of "moderate" or "neutral" attitudes.

REFERENCES

1. DIAB, L. N. Studies in social attitudes: I. Variations in latitudes of acceptance and rejection as a function of varying position on a controversial social issue. *J. Soc. Psychol.*, 67, 283-295.
2. HOVLAND, C. I., HARVEY, O. J., & SHERIF, M. Assimilation and contrast effects in reaction to communication and attitude change. *J. Abn. & Soc. Psychol.*, 1957, 55, 244-252.
3. MAXWELL, A. E. *Analysing Qualitative Data*. London: Methuen, 1961.
4. SHERIF, M. Some needed concepts in the study of social attitudes. In J. G. Peatman & E. L. Hartley (Eds.), *Festschrift for Gardner Murphy*. New York: Harper, 1960.
5. SHERIF, M., & HOVLAND, C. I. *Social Judgement*. New Haven: Yale Univ. Press, 1961.

Department of Psychology
American University of Beirut
Beirut, Lebanon

STUDIES IN SOCIAL ATTITUDES: III. ATTITUDE
ASSESSMENT THROUGH THE SEMANTIC-
DIFFERENTIAL TECHNIQUE*¹

Department of Psychology, American University of Beirut

LUTFY N. DIAB

A. INTRODUCTION

In a recent study reported by the author (1), it was found that subjects with a moderate stand on the issue of Arab unity, as measured by the Sherif-Hovland technique (6, 7), rejected significantly more items than they accepted from a total of nine items (A to I) representing the evaluative (good-bad) dimension of Arab unity. Such a finding was expected only for subjects with an extreme positive or negative stand toward the issue in question, but not for those with a moderate stand. In a follow-up study by the same author (2), it was also found that subjects with a moderate stand on the Arab-unity issue showed, in their reactions to various kinds of mass-communication media, selective behavior similar to that of the extreme anti-Arab unity subjects. Thus, doubt was then expressed as to the neutrality of so-called moderate subjects, as measured by the Sherif-Hovland technique (2, 7).

The present study attempts to find explanations for the foregoing discrepancies, by utilizing the semantic-differential technique which, according to its authors (5), should enable one to get a more-complete picture of the meaning of "Arab unity" than is obtained through the evaluative dimension alone. Furthermore, the use of more than one evaluative scale to represent the evaluative dimension should give a more-exact estimate of the subjects' attitude scores. Finally, the use of a variety of semantic-differential scales, representing various dimensions of meaning, should clarify the behavior of subjects with extreme positive or negative stands toward the issue of "Arab unity."

According to Osgood, Suci, and Tannenbaum, "... to index attitude we would use sets of scales which have high loadings on the evaluative factor across concepts generally and negligible loadings on other factors . . ." (5, p. 191). Thus, to Osgood *et al.* attitude is simply identified as the evaluative

* Received in the Editorial Office, Provincetown, Massachusetts, on May 25, 1964.
Copyright, 1965, by The Journal Press.

¹ Thanks are extended to Miss Suad Kaylani for her assistance in tabulating the data.

dimension of the total semantic space. They also report that in all of the factor analyses done by them the evaluative dimension accounted "for the largest proportion of the total variance" (5, p. 190). Furthermore, in measuring attitude, they state ". . . we usually include a considerable number of scales representing other factors—this is done both to obscure somewhat the purpose of the measurement and to provide additional information on the meaning of the concept as a whole, aside from the attitude toward it" (5, p. 191). Thus, our present choice of the semantic differential *in toto* as a measuring instrument was based on the evidence presented by Osgood *et al.*, which shows that the differential provides "a richer picture of the meaning of the attitude object than just the evaluative dimension . . ." (5, p. 195). Osgood *et al.* conclude that "by combining judgments derived from scales representing other dimensions with those derived from the evaluative factor alone, additional information can be obtained and prediction presumably improved" (5, p. 198). In the light of accumulated evidence, Osgood and his colleagues report the cross-cultural and cross-linguistic existence of the same three independent factors or dimensions that were found with American subjects (3). These factors are designated the *evaluative*, *potency*, and *activity* factors of meaning. Our present interest was in selecting scales with high loadings on the evaluative factor for the Arabic language and in the use of such scales (consisting of pairs of adjectival opposites) as a basis for judging the concept "Arab unity." Such scales, representing the evaluative factor, as well as other scales representing other factors in the semantic space for our Arab subjects, were available in a mimeographed report by Osgood, Archer, and Miron (4); hence the present study will make use of a number of semantic-differential scales, representing the evaluative as well as other factors of meaning, in an attempt to clarify some of the unexpected findings concerning moderate subjects (1, 2) and to throw more light on the meaning of "Arab unity" to extreme pro-Arab unity as well as to extreme anti-Arab unity subjects.

B. METHOD

1. Subjects

The same sample used earlier by the author (1, 2) was also utilized in this study. It consisted of 260 undergraduate Arab students at the American University of Beirut (198 males and 62 females) with a mean age of 20½ years for the total sample. As to religion, 136 were Moslems, 111 were Christians, and the rest indicated no religion or belonged to the Druze sect. In terms of nationality, 124 were Lebanese, 94 were Jordanians or Palestinians, and the rest were mainly Syrians, Sudanese, and Iraqis.

2. Procedure

In addition to information obtained previously and mentioned in earlier studies (1, 2), the questionnaire administered anonymously to the subjects included also a series of 13 different scales against which the concept of "Arab unity" was to be judged, following exactly Osgood's procedure and instructions (4, 5). Four of these scales have very high loadings on the "evaluative" factor and negligible loadings on the other factors (4). These were the good-bad, beneficial-injurious, safe-dangerous, and great-contemptible scales. Another set of three scales were chosen (4) as representative of the "potency" factor. These were the slow-fast, strong-weak, and rational-emotional scales. The remaining six scales were selected in such a way as to represent the "activity" and other unidentified factors or dimensions, particularly those that appeared to the author to be most relevant to the issue of Arab unity.

C. RESULTS

The use of the semantic-differential technique was intended, among other things, to provide further clarifications for some of the unexpected findings concerning moderate subjects, reached in earlier studies (1, 2). As stated in the introduction, subjects with a moderate stand towards Arab unity (those checking Positions D, E, or F as "most acceptable" from a good-bad continuum of nine statements—A to I) were found to reject a significantly greater number of statements than they accepted. Such an unexpected finding for moderate subjects was previously accounted for in terms of the strong commitment to the moderate stand exhibited by the majority of moderate subjects (1). In addition to this explanation, the use of several semantic-differential evaluation scales has helped to shed more light on the behavior of moderate subjects. As shown in Table 1, our so-called moderate subjects, as measured by the Sherif-Hovland technique (1, 6, 7), do not seem to be moderate, as measured by their mean evaluative ratings on four semantic-differential evaluative scales. In fact, moderate subjects at Position D have a mean evaluative rating of 2.28 on a seven-point scale, which means that they feel that Arab unity is mostly "quite good, beneficial, safe, and great." On the other hand, moderate subjects at Position F have a mean evaluative rating of 5.96, which indicates that Arab unity is, from their point of view, "quite bad, injurious, dangerous, and contemptible." Only those moderate subjects at Position E turned out to be near the point of neutrality, with a mean evaluative rating of 3.81, indicating that Arab unity is to them "equally good and bad." Thus, it appears from these results that more than half of our so-called moderate subjects (about 62 per cent) are not moderates, but

TABLE 1
MEAN RATINGS OF THE CONCEPT "ARAB UNITY" ON 13 SEMANTIC-DIFFERENTIAL SCALES

Dimension	Position checked "Most Acceptable" on evaluative dimension								
	A (N = 99)	B (N = 60)	C (N = 19)	D (N = 25)	E (N = 20)	F (N = 8)	G (N = 6)	H (N = 7)	I (N = 11)
Evaluative scales									
1. Good-bad	1.12	1.35	1.68	2.04	3.45	5.75	6.00	5.71	6.18
2. Beneficial-injurious	1.19	1.45	2.00	1.92	3.65	6.12	5.00	6.43	5.82
3. Safe-dangerous	1.38	2.10	1.79	3.00	4.30	6.12	5.33	6.71	5.82
4. Great-contemptible	1.34	1.58	2.05	2.16	3.85	5.87	5.33	5.85	5.73
Mean	1.26	1.62	1.88	2.28	3.81	5.96	5.41	6.17	5.89
Potency scales									
5. Strong-weak	1.43	1.81	2.10	2.68	4.25	5.87	5.16	6.71	5.55
6. Rational-emotional	2.96	3.43	4.47	4.72	5.20	6.87	4.33	7.00	6.00
7. Fast-slow	4.04	4.30	4.42	5.88	5.85	4.75	6.33	5.86	7.00
Mean	2.10	3.18	3.66	4.43	5.10	5.83	5.27	6.52	6.18
Other scales									
8. Near-far	2.71	3.36	3.89	5.28	5.22	6.87	5.50	6.85	6.18
9. Active-passive	1.58	2.11	2.79	3.48	3.70	6.62	5.00	5.14	5.45
10. Simple-difficult	4.73	5.43	5.57	5.96	6.00	6.87	5.66	6.71	6.73
11. Real-imaginary	1.41	1.80	2.58	3.92	3.90	6.62	5.33	6.71	6.00
12. Possible-impossible*	1.51	1.63	2.16	3.28	4.50	6.75	6.52	5.28	6.36
13. Predictable-unpredictable*	1.90	2.65	2.84	4.08	4.75	6.50	5.83	5.43	5.64

* With the exception of the last two scales, which were added by the author because of their relevance to the concept in question, all other scales were taken from factor analyses on a variety of scales for the Arabic language and reported by Osgood, Archer, and Miron (4).

are, instead, either for or against Arab unity [see Table 1]. In other words, more than half of our moderate subjects should have been classified, in the light of this additional information provided by the use of four evaluative scales of the semantic-differential technique, as either extreme pro-Arab unity or extreme anti-Arab unity subjects rather than moderate subjects.

Furthermore, as shown in Table 1, about 53 per cent of our so-called moderate subjects have mean potency ratings similar to those of extreme anti-Arab unity subjects, indicating their feelings that Arab unity is to them "quite weak, emotional, and slow." In addition, moderates have similar mean ratings as anti-Arab unity subjects on such scales as "near—far," "simple—difficult," and "predictable—unpredictable." Thus, on the basis of these findings on our so-called moderate subjects, it is clear that the meaning of "Arab unity" to them has a lot in common with what this concept means to extreme anti-Arab unity subjects.

Finally, this study was also intended to throw more light on the meaning of "Arab unity" to subjects with extreme positive as well as negative stands towards the issue. Thus, as shown in Table 2, the sample was divided into seven groups of subjects, depending on the rating they gave for the concept "Arab unity" on the seven-point good-bad scale. For example, there were 151 subjects who gave a rating of 1.00 on the good-bad scale, indicating their feeling that Arab unity is "very good"; while, on the other hand, there were 22 subjects who gave a rating of 7.00 on the same scale, indicating their feeling that Arab unity is "very bad." Our results indicate that those subjects who evaluate Arab unity "quite" or "very" highly (i.e., subjects with an extreme positive stand or attitude towards Arab unity) do not seem to have ratings of equal magnitudes on scales representing dimensions or factors other than the evaluative factor. As Table 2 shows, these same pro-Arab unity subjects feel that Arab unity is either slightly potent or somewhere near the point of neutrality as far as potency is concerned. Furthermore, they also feel that Arab unity is "quite or slightly active, real, possible, and predictable" but also "slightly far and quite or slightly difficult" to achieve [see Table 2].

As to subjects with an extreme negative stand toward the Arab unity issue (i.e., those having mean evaluative ratings of about 6.00 or above and thus indicating that Arab unity is to them "quite bad" or "very bad"), these subjects appear to have ratings of equal magnitudes on scales representing factors other than the evaluative factor. Thus, these same anti-Arab unity subjects also feel that Arab unity is "quite or very low in potency" as well as "quite or very far, passive, difficult, imaginary, impossible, and unpredictable" [see Table 2]. In the case of so-called moderate subjects (i.e., those having mean

TABLE 2
MEAN GOOD-BAD RATINGS OF THE CONCEPT "ARAB UNITY" ON 13 SEMANTIC-DIFFERENTIAL SCALES

Scale	Good:			equally (<i>N</i> = 12)	Bad:		
	very (<i>N</i> = 151)	quite (<i>N</i> = 56)	slightly (<i>N</i> = 7)		slightly (<i>N</i> = 2)	quite (<i>N</i> = 7)	very (<i>N</i> = 22)
Evaluative scales							
1. Good-bad	1.00	2.00	3.00	4.00	5.00	6.00	7.00
2. Beneficial-injurious	1.05	2.00	3.14	3.50	4.50	5.57	6.59
3. Safe-dangerous	1.56	2.87	3.00	4.25	5.50	6.00	6.36
4. Great-contemptible	1.35	2.45	2.57	3.67	3.00	4.71	6.64
Mean	1.24	2.33	2.93	3.85	4.50	5.57	6.65
Potency scales							
5. Strong-weak	1.48	2.68	2.86	3.50	6.50	6.00	6.59
6. Rational-emotional	3.02	4.38	4.86	5.83	5.50	5.86	6.18
7. Fast-slow	4.10	5.23	5.14	5.67	6.50	4.71	6.86
Mean	2.87	4.10	4.29	5.00	6.17	5.52	6.54
Other scales							
8. Near-far	2.93	4.71	4.14	4.67	6.50	6.57	6.59
9. Active-passive	1.73	3.14	3.00	3.83	4.00	6.57	5.73
10. Simple-difficult	5.05	5.59	5.43	6.17	7.00	6.28	6.59
11. Real-imaginary	1.56	2.96	3.43	4.17	6.00	6.14	6.50
12. Possible-impossible*	1.54	2.78	3.29	4.42	4.50	6.14	6.77
13. Predictable-unpredictable*	2.16	3.55	3.86	4.58	4.50	6.00	6.09

* With the exception of the last two scales, which were added by the author because of their relevance to the concept in question, all other scales were taken from factor analyses of a variety of scales for the Arabic language and reported by Osgood, Archer, and Miron (4).

evaluative ratings of about 4.00 and thus feeling that Arab unity is about "equally good and bad"), it appears that they judge "Arab unity" on the potency factor in a manner similar to that shown by the anti-Arab unity subjects, as indicated by their mean potency rating of 5.00 across all three potency scales [see Table 2]. Also, moderate subjects on the evaluative factor judge "Arab unity," similar to anti-Arab unity subjects, as being "slightly far, impossible and unpredictable" as well as "quite difficult" to achieve.

D. DISCUSSION

The results of this study demonstrated that the use of various semantic-differential scales (representative of the evaluative, potency, and other factors of meaning), in contrast to the use of scales representative of the evaluative factor alone, results in a more-complete picture about the attitude object and does help to improve prediction. As stated by Osgood *et al.* (5, p. 198), "two people may have identical *attitudes* toward a concept (as determined by allocation to the evaluative dimension alone), and yet have quite different meanings of the concept (as determined by the profiles as wholes)."

The results presented in Table 1 show a comparison of the information that can be obtained from using scales representative of various factors of meaning to that which was obtained on the same subjects in an earlier study, using the Sherif-Hovland technique (1). About 62 per cent of the subjects that were considered previously as moderates (at positions, D, E, and F) on the evaluative dimension turned out, as shown by their mean evaluative ratings [see Table 1], to be either strongly for or strongly against Arab unity rather than moderates. Thus, if the majority of so-called moderates have extreme rather than neutral stands, as revealed by the foregoing results, then it is reasonable that they would reject more statements than they accept (1). It may very well be that for our Arab subjects a truly neutral or moderate stand is impossible to find. In other words, it may mean that those subjects who consider themselves to be moderates are not really expressing their true feelings, and that the use of less-direct techniques of questioning may help to reveal this tendency.

On one hand, the Sherif-Hovland technique makes use of a series of sentences or statements (A to I) that vary all the way from "the achievement of Arab unity is absolutely essential from all angles in the Arab people's interests" to the extreme opposite of this statement. And, among other things, the subject is simply asked to indicate the statement "most acceptable" to him. On the basis of classifying subjects in terms of statements "most acceptable" to them, our moderate subjects were those checking Positions D,

E, and F. On the other hand, the semantic-differential technique uses scales of adjectival opposites instead of sentences or statements and such usage may result in less conscious control on the part of the subject and a greater readiness to respond with his true feeling about the concept. Thus, our results have shown that moderate subjects at Position D and at Position F (as determined by the most-acceptable statement to them from a series of nine statements—A to I) were, respectively, pro-Arab unity and anti-Arab unity (as measured by their mean ratings on four evaluative scales of the semantic differential), rather than moderates. In other words, it seems that the use of scales of adjectival opposites results in a more-correct estimate of one's stand than when such stand is determined by the sentence or the statement "most acceptable" to the individual.

Not only did the so-called moderate subjects reject significantly more statements than they accepted on the evaluative dimension (1), but they also showed selectivity in mass-communication media (i.e., newspapers and radio stations) similar to that shown by anti-Arab unity subjects (2). Again, such similarity to anti-Arab unity subjects compels us to doubt the neutrality of these so-called moderates, as determined by the statement "most acceptable" to them. Additional information confirming our doubts is provided by the general view that these subjects have of Arab unity, as indicated by their mean ratings on a number of scales representing potency and other factors of meaning [see Table 1]. Thus, our results have shown that these subjects view Arab unity as "quite weak, emotional, slow, far, and difficult." In short, it seems that the meaning of the concept "Arab unity" to so-called moderate subjects (at Positions D, E, and F) is similar in a number of respects to the meaning that the concept has to anti-Arab unity subjects (at Positions G, H, and I).

Finally, subjects were classified in terms of the rating they gave for "Arab unity" on the good-bad scale, as representative of the evaluative factor, in order to find out the differences in meaning of the concept for various groups of subjects having different attitudes. This classification of our sample was followed because we are interested in finding out if differences in evaluation of the attitude object are necessarily associated with similar systematic differences in the other factors of meaning. Furthermore, this procedure was followed because (on the one hand) we have here in the concept of "Arab unity" a controversial issue that is highly ego involving for our subjects and (on the other hand) because the evaluative factor accounts for approximately 54 per cent of the total variance (4). Thus, as may be seen from Table 2, our results show that subjects who evaluated "Arab unity" favorably, as shown by

their mean evaluative ratings, had near-neutral mean potency ratings and felt also that Arab unity is "slightly far as well as quite or slightly difficult." In other words, the same subjects who feel that Arab unity is "very good" or "quite good" also feel that such unity is "not so near," "quite difficult," but "quite or slightly active, real, possible, and predictable." This picture that subjects with an extreme-positive stand towards Arab unity have could not have been known or predicted from their positive evaluative stand alone. Thus, the use of scales representative of factors other than the evaluative factor alone, enabled us to get a much more-complete picture about the meaning of the concept of Arab unity to our extreme pro-Arab unity subjects.

Now, what about our anti-Arab unity subjects? As shown in Table 2, our results show that subjects who evaluated "Arab unity" unfavorably, as shown by their mean evaluative ratings, also had mean ratings indicating low potency, in addition to their other expressed feelings that Arab unity is "very far," "quite or very passive," "very difficult," "very imaginary and impossible," as well as "quite unpredictable." It seems that subjects who feel that Arab unity is "very bad" or "quite bad" have a tendency, regardless of which scale of adjectival opposites is presented to them, to check the "apparently" unfavorable end of the scale as best representative of their feeling. Thus, because Arab unity is "very bad," it is also "very passive, difficult, imaginary, impossible, and even unpredictable." One may notice here the contradictions involved in believing that Arab unity is "impossible" but at the same time "unpredictable." Also, similar inconsistencies appear in their feelings that Arab unity is "weak, passive, and imaginary," yet at the same time "injurious and dangerous." All kinds of explanations can be offered here to account for these contradictions and inconsistencies; but, at the same time, it must be pointed out that from the point of view of the anti-Arab unity subjects themselves these contradictions are perhaps neither recognized nor admitted as such. If Arab unity is "imaginary and impossible," then one should be able to "predict" that it will not materialize; however, being "predictable" may have had a "favorable" flavor about it, thus implying actualization of unity and, therefore, making our anti-Arab unity subjects feel that Arab unity is "unpredictable." That Arab unity is "dangerous and injurious" to our anti-Arab unity subjects throws further light on our previous finding showing that these subjects, much more so than pro-Arab unity subjects, were highly selective in the kinds of mass-communication media to which they expose themselves, the selectivity being in the direction of media expressing their own points of view about the issue in question (2). It is only reasonable to explain this "avoidance" behavior of anti-Arab unity subjects

as related to their conception of Arab unity as "dangerous and injurious." Thus, on the whole, it appears that those subjects who evaluate Arab unity unfavorably tend also to pick the "apparently" unfavorable end of any scale (be it representative of potency, activity, or other factors) as best expressing their feeling.

As to subjects with neutral mean evaluative ratings, our results show that they judge Arab unity as "quite slow and emotional" and, in a manner similar to anti-Arab unity subjects, "slightly far, impossible, and unpredictable" as well as "quite difficult" [see Table 2]. Again, the tendency on the part of subjects who evaluate Arab unity as "equally good and bad" to view Arab unity in terms similar to anti-Arab unity subjects gives further explanation to the similarity in selectivity of mass-communication media previously between moderate and anti-Arab unity subjects (2).

In concluding this discussion, it seems to the author that accuracy in measurement and consequently in prediction can be considerably improved if a combination of the Sherif-Hovland technique (7) and the Osgood semantic-differential procedure (5) is used. On one hand, the Sherif-Hovland technique provides not simply a "point" or "score" but a "latitude of acceptance" and a "latitude of rejection" on any continuum or scale, thus increasing our information about the subject's stand on an issue and reflecting more accurately the realities of life situations. On the other hand, the semantic-differential procedure affords richer information about the meaning of the attitude object through the use of scales found to be representative of evaluation, potency, activity, and other factors. In addition, the semantic-differential technique is a simple "procedure in which a single stimulus is judged successively against a series of different scales" (3, p. 26) representative of the various independent factors of meaning.

About six scales representative of the various independent factors of meaning (i.e., evaluation, potency, etc.) and appropriate to the concept in question are presented to the subjects in the regular semantic-differential rating form. However, for each of the scales, subjects are asked to indicate the position "most acceptable" as well as other positions "also acceptable" to them. The individual's score on each scale would be the mean value of his latitude of acceptance: i.e., his mean acceptable position. This modification should give us more information about the meaning of a concept than either of the above techniques alone and consequently should help improve prediction considerably. Such a procedure is worth investigating more thoroughly in future research.

E. SUMMARY

The purpose of the present study was to use the semantic-differential technique in an attempt to throw more light on the meaning of "Arab unity" to subjects with various stands and to clarify some of the unexpected findings reached in earlier studies (1, 2).

The same sample used in earlier studies (1, 2), consisting of 260 undergraduate Arab students at the American University of Beirut, was utilized in this study. In addition to information obtained previously (1, 2), the subjects were given a series of 13 different scales against each of which the concept of "Arab unity" was to be judged successively, according to the semantic-differential procedure.

In general, the results of this study show that additional information obtained from scales representative of evaluative as well as other factors on the meaning of the concept "Arab unity" help to clarify some of the unexpected findings of previous studies (1, 2). Thus, subjects considered previously to be moderates turned out to be mostly for or against Arab unity, as shown by their mean evaluative ratings. Consequently, the behavior of so-called moderates in rejecting significantly more statements than they accept (1) and in exhibiting selectivity in mass-communication media similar to anti-Arab unity subjects (2) are explained more fully in the light of additional information obtained from the use of the semantic-differential technique. Also, the meaning of "Arab unity" to pro-Arab unity and anti-Arab unity subjects is more fully understood as a result of information provided by results obtained from the semantic differential.

Finally, a modified technique is suggested by the author in which scales representative of the total semantic differential are presented to the subject and, in line with the Sherif-Hovland technique, latitudes of acceptance are obtained on each of the semantic-differential scales. Such a procedure appears to help in improving prediction and is worth further investigation.

REFERENCES

1. DIAB, L. N. Studies in social attitudes: I. Variations in latitudes of acceptance and rejection as a function of varying positions on a controversial social issue. *J. Soc. Psychol.*, **67**, 283-295.
2. ———. Studies in social attitudes: II. Selectivity in mass communication media as a function of attitude-medium discrepancy. *J. Soc. Psychol.*, **67**, 297-302.
3. OSGOOD, C. E. Studies on the generality of affective meaning systems. *Amer. Psychologist*, 1962, **17**, 10-28.
4. OSGOOD, C. E., ARCHER, W. K., & MIRON, M. S. The Cross-Cultural Generality of Meaning Systems: A Progress Report (mimeographed). University of Illinois, Urbana, Illinois, 1963.

5. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. The Measurement of Meaning. Urbana: Univ. Illinois Press, 1957.
6. SHERIF, M. Some needed concepts in the study of social attitudes. In J. G. Peatman & E. L. Hartley (Eds.), *Festschrift for Gardner Murphy*. New York: Harper, 1960.
7. SHERIF, M., & HOVLAND, C. I. Social Judgment. New Haven: Yale Univ. Press, 1961.

Department of Psychology
American University of Beirut
Beirut, Lebanon

EXPERIMENTAL SOCIAL PSYCHOLOGY AND ITS NEIGHBORS*¹

Department of Psychology, University of Washington

EZRA STOTLAND

A. INTRODUCTION

Frequently social psychologists find themselves the targets of a variety of semirhetorical questions: How can there be a valid scientific theory in social psychology? How can there be any scientific laws in social psychology? and so forth. Sometimes social psychologists attempt to answer these questions, sometimes not. Often, social psychologists ignore the questions in public; but, when the critics have gone home, they think about them. This paper is the result of one social psychologist's attempt to come to terms with a few questions like those posed. Three questions are considered. Each concerns the relationship between social psychology (especially laboratory social psychology) and one of its neighbors: other types of experimental psychology, common sense, and field studies.

B. QUESTIONS

1. *Can There Be a "Social" Psychology Separate from or Independent of the Rest of Psychology?*

Can social psychology be "reduced" or eliminated by being absorbed into the rest of psychology? This question can be answered only by making an arbitrary but not meaningless distinction between two levels of discourse in social psychology (with full recognition of the existence of many shades of grey between the two levels). One level concerns the individual reacting to a situation that includes other people or symbols of other people. Any laws that are developed for this level are laws in which the individual's reaction is the dependent variable. Let us call such laws individual-social laws. At another level, dependent variables concern the relationships between or among persons: e.g., leadership, a role structure, a group process, etc. Let us call these laws group laws.

Individual-social laws bear a close family resemblance to the laws in general

* Received in the Editorial Office, Provincetown, Massachusetts, on June 1, 1964. Copyright, 1965, by The Journal Press.

¹ This paper was written while the author was a USPHS Fellow at the University of London. He is greatly indebted to Judson Mills and Nathan Maccoby for their critical readings of early drafts.

psychology. They can and have been translated into simple S-R laws, laws of perception, etc. Our question is, "Are individual-social laws different in any fundamental way from other laws in psychology?"

My answer is that they are *not* different in a fundamental way, but that the values of certain parameters in the individual-social laws have to be considered in comparison with the values of those parameters in the laws of general psychology. The general laws and the individual-social laws may vary systematically with variations in the values of the parameters. For example, suppose a parameter to be ambiguity of the stimulus situation. In general, but not always, social situations are more ambiguous than nonsocial situations. It may be that certain laws relating motivation to perception take one form in clearly defined situations and another form in more ambiguously defined ones. If so, one would expect that the application of these laws to social situations would necessitate consideration of the generally greater ambiguity of social situations.

My argument is not for the reduction of social-individual laws to general laws; it is for the "reduction" of the differences between a social-individual hypothesis and what usually is called a general psychological hypothesis to a number of differences in parameter values (e.g., ambiguity, type of motivation, degree of similarity of subject to the source of stimulation) and for a systematic consideration of whether or not the differences in parameter values lead to differences in the laws.

The foregoing approach has several advantages. First, it implies study of variations along each parameter while holding the type of law constant. Not all social situations are ambiguous; not all nonsocial situations are clearly defined. Social situations would be varied in ambiguity; so would the nonsocial. If the universe is as uniform as generally it is thought to be, the variations in the degree of ambiguity should have the same effect on the form of the laws in individual-social conditions as in the nonsocial conditions. Thus, the first potential advantage is parsimony.

Second, the second-order law relating the parameter value to the first-order law would receive increased empirical support from the data coming from the social and nonsocial realms; thus the confidence that could be placed in the law within each realm would be increased by the knowledge that it derives support in more than one realm.

Third, if variations of parameter values in individual-social and nonsocial realms lead to variations in the form of the laws, greater precision of prediction can be attained. Because social situations or nonsocial situations are not equally ambiguous or equally well structured, consideration of the degree

of ambiguity within the social or within the nonsocial realms should lead to more exact predictions than otherwise. Attempts to place the individual-social laws and nonsocial laws in the same rubric can have heuristic value, suggesting new parameters in both areas or suggesting new types of laws. It is important to note that, in this way, the general psychologist derives useful suggestions from his social-psychologist neighbor.

Fourth, if the relationship between parameter values and the form of the laws should prove as uniform as suggested, any explanations of the laws should apply equally well to the social and nonsocial situations. By explanation of the laws, I mean those mediating processes (within the individual or in his history) that are assumed to be necessary for the lawful relationships to obtain. Again, greater parsimony is achieved.

At the second level of law, the group law, there is no analogous type in general psychology. There is no way of saying meaningfully that, after a certain amount of cooperation between a person and a machine, a mutual attraction develops. In other words, the laws in general psychology cannot be treated in terms of variations on a parameter. These group laws, as statements of regular empirical relationships, cannot be special cases of general laws that apply also to nonsocial situations. Despite the foregoing, the separation on the level of empirical laws should not lead one to the conclusion that these laws are floating in scientific space with no anchors. The anchors become evident when an attempt is made to explain the laws. In this case, the explanation of the laws may consist of a set of individual-social laws or individual nonsocial laws.

For example, we might attempt to explain why two or more people who have successfully cooperated as part of a group develop a mutual attraction: i.e., that group cohesiveness rises. We might invoke some law, such as the following: If an individual is rewarded regularly by some entity (either social or nonsocial) he seeks to be in the presence of that entity. This law is then applied to each relationship of one group member to another (i.e., a series of one-to-one relationships). A logical outcome of the application of the law to the series of relationships is the empirical law relating successful cooperation to group cohesiveness. The group law is not eliminated. It is retained in its original form, but it has been explained and (through the explanation) has been related to the total body of psychological knowledge; thus, indirectly, to all knowledge.

This process of explaining laws can be attempted, of course, at all levels of law: the individual-social, the nonsocial. Laws might be explained by cognitive processes, by neural processes, etc. The process does not explain

away the laws. The laws remain, neither as simply arbitrary facts of nature nor as "regularly occurring accidents," but remain as part of a general body of laws.

But why retain the laws if they can be explained in principle or in fact? There are a variety of reasons. First, the exact theoretical formulation or the empirical basis of the explanation of the laws may not have been attained. To say that such explanations are possible in principle is not sufficient with regard to any particular law. It is also necessary to point out how the explanation applies (*cf.* 4).

Another reason for retaining the laws applies after the explanations have been discovered. This reason is that the derivation of the law from its explanation may be so complex that the law is a useful shorthand. It is impractical and pointless to go through the process of logically deriving the law relating successful cooperation and mutual attraction over and over again for each case to which the law applies. The issue is more important if the law is to be used as the basis of explaining laws at a still higher level of abstraction. Suppose that a researcher is interested in developing a law concerning the stability of large institutions, such as factories or other firms. He may need to assume the validity of the group laws in attempting to explain laws at the institutional level. The researcher cannot start down the trail of reductionism that leads to abandoning his institutional laws and to formulating everything in terms of subatomic particles. No nonnuclear scientist, no general psychologist would have a *raison-d'être*.

2. *How Does Social Psychology Go Beyond Common Sense?*

Unfortunately, there is much evidence for the accusation implicit in the question. Social psychologists have tested experimentally theories such as "a person will do something if it gets him something he wants and if it doesn't cost too much"; "if someone agrees with you, you like him more"; and "if you frustrate someone, he will become angry." In each case, as in all cases of common-sense propositions, no amount of laboratory testing will refute the fact that at least one and probably many instances consistent with the proposition can be found in the real-life experience of social psychologists, other psychologists, and just plain people. If the laboratory test does not support one of the propositions, it is the test that frequently is rejected, not the proposition. The test can be rejected on grounds of artificiality, superficiality, erroneous assumptions, or (for the "know-nothing") on grounds of "so what?" Because so many assumptions are involved even in the simplest of experiments, it is just as easy to reject the assumptions and the test as it is to reject the hypothesis.

In defense of social psychologists, one often argues that one man's common sense is another man's puzzle. But this fact does not contradict the point that, in some cases, there can be no question that the propositions or hypotheses are of a common-sense nature.

The germ of a defense for social psychology emerges when contradictory common-sense propositions are placed side by side. The scientific question, however, should not be formulated in terms of which proposition is true and which is not. As indicated, from real life one can find instances consistent with both propositions. The scientific question is, "Under what conditions is one proposition true; under what conditions is the other true?" The most useful experimental design is one in which those conditions are determined. Often one requires an experimental design in which both propositions are tested under conditions varied systematically. If only one proposition is tested, all one can conclude is that under certain specified conditions the proposition tested is valid. One can say nothing about the truth or falsity of the other proposition. The scientific question can be answered only if the sufficient conditions for each of the propositions have been ascertained.

The approach of finding the sufficient and, in some cases, the necessary conditions for a proposition to be true does not apply only to cases of contradictory common-sense propositions. It can be made to apply to some propositions that do not appear to have any commonly believed contradictions. For example, the author does not know of any common American proverb that contradicts the idea that "we like people who help us," yet question can be raised as to whether this proposition is universal. A researcher might ask "Under what conditions is it false?" Perhaps the perceived motivations of the helper are relevant.

The same approach can be applied to common-sense propositions, such as maximizing gains and minimizing losses, provided some independent criterion is established for "gains" and "losses." "Under what conditions do people *not* maximize gains and minimize losses, even when they clearly and accurately perceive the consequences of their actions?" Dostoevsky contended that people do not and should not always act in such a rational way.

There is another way of handling the common-sense argument. In real life, the sufficient conditions for a process often are intertwined with nonrelevant conditions to such an extent that it is impossible to separate one from the other. For example, if the proposition under scrutiny is that "we like someone who agrees with us," what are the sufficient conditions for liking to occur? Is it sufficient that we are aware of the fact that the person agrees? Is it sufficient that he communicates this agreement to us? Is it sufficient that we observe him in the process of coming to agree with us? The researcher

must attempt to discover which one or more of these conditions are sufficient.

The foregoing is counter to the advice sometimes given [see French (2)] that we "start strong" in the laboratory, that we attempt to administer the independent variable by using a single induction that contains as many different manifestations or aspects of the independent variable as possible: e.g., several dimensions of status. If the phenomenon has been observed in at least one instance in real life, no new knowledge is gained by observing the phenomenon again in the laboratory; and "starting strong" sometimes may accomplish only that. By "starting weak" (i.e., by inducing the independent variable in one precisely defined way), we may come closer to discovering the conditions sufficient for the dependent phenomena to occur. Of course, we run the risk of having bet on the wrong condition; but that risk is what makes research fascinating.

The foregoing argument implies that if the hypothesis in the laboratory concerns a process that does not fit into common sense one may find it helpful to "start strong" at the beginning of a program of research. The first step in such a program is to establish that the phenomena does in fact occur—a finding that should surprise everyone but the researcher and some astute laymen. It is at this point that the laboratory is at its most productive and interesting.

3. *How Can We Generalize from Laboratory To Real Life?*

To answer this question, it is necessary to subdivide it into several smaller questions. First, "What sorts of real-life questions is it possible to answer in the laboratory?" There are at least two types of questions that cannot be resolved in the laboratory. One often has been pointed out. It is neither ethical nor practical to deal with certain variables in the laboratory. Such variables are suicide, extremely stressful conditions, and extreme aggression. Sometimes it is possible in the laboratory to deal usefully with watered-down versions of the foregoing variables, but one should not mistake a watered-down version for the "real thing." Factors that function in the laboratory might not function in the real-life condition. One must be careful in making extrapolations from the laboratory to real-life situations or, for that matter, vice versa.

A second type of question that cannot be resolved definitively in the laboratory is a chicken-and-egg puzzle that often grows out of field research or even correlational survey research, although some researchers [e.g., Festinger (1)] have stated that the laboratory can resolve the puzzle. A question often asked is, "Does A cause B or does B cause A?" This type of question *cannot* be resolved definitively in the laboratory.

In the first place, because of the way the question is sometimes formulated there is no reason *not* to believe that both alternatives are true: i.e., that A and B influence one another. Of course, in many cases, such mutual influence is simply not possible; but, in the cases for which the question does arise, it is evident that both directions of causation are considered possible. For mutual influence to occur, there must be some time lag in the effects. If A caused B instantaneously, and B caused A instantaneously, A and B would fuse into a sort of steady-state unity that could be considered as one entity. For most social-psychological phenomena there is a time lag.

Let us assume that it is possible for A to cause B and for B to cause A. If, in the laboratory, we demonstrate that A causes B, that fact does not contradict the proposition that B causes A. At this point, the laboratory has accomplished little. Furthermore, if it is found that A does not cause B in the laboratory, it still might do so in real life. The finding in the laboratory may be interesting, but it is ambiguous because there are many factors that might account for the failure. If the question is reformulated as, "Does A cause B more than B causes A?" the answer is to be found in the field situation and not in the laboratory, because the factors enhancing one direction of causation rather than the other might be present only in the real-life situation.

What, then, can the laboratory accomplish in chicken-and-egg problems? The answers basically are the same as those given to the questions about common sense. When there is serious doubt as to whether A leads to B or B leads to A or both, the laboratory can establish that, under certain circumstances, one or both of the sequences occurs. This demonstration is the laboratory's most useful function.

Note that the negative does not apply. The laboratory cannot prove that a given sequence does *not* occur in real life, but when there is no doubt about the occurrence of a sequence in real life, the laboratory might provide evidence with regard to the conditions sufficient for one of the sequences.

The essence of the argument is that neither the laboratory findings nor the field studies can refute one another. If the results are inconsistent, the problem is to discover what parameters account for the difference. After all, behavior in the laboratory is just as real as behavior anywhere else. The problem is to interpret theoretically the findings—not to have them clash [*cf.* Hovland (3)].

What are the bases for generalization from the laboratory to real life? Of course, there are the useful warnings about populations of people and of situations. However, such warnings are of no use unless we know some way of taking them into consideration. A first approximation to an answer derives from an examination of the theoretical variables that are assumed

to have been actualized in the laboratory. Presumably, we should be able to generalize to other instances of these theoretical variables.

Parallel statements can be made for the population of subjects. The rub lies, however, in the fact that any concrete event or sequence can be construed to be an instance of an almost infinite variety of abstractions. The question becomes, "What grounds can the researcher have for interpreting his experiment in the way that he does?" The answer lies in either one or both of two directions. First, the researcher may have data that can be construed as reflecting other instances of the same phenomena. These instances may be the same phenomena or the same type of phenomena. For example, there may be instances of leadership roles showing a given effect on the role-player or instances of roles other than leadership roles having the same effect. The rub in this type of answer is the excessive ease and simplicity of the construing. However, if enough instances can be cited the easiest and simplest construction should become apparent. These instances should, of course, come from as wide a variety of laboratory and field studies as possible. (This type of argument also applies to the relationships among field studies.)

In some instances, the real-life "replication" of a laboratory situation may not have occurred anywhere in society or may not have been recorded; yet social changes and continued studies in the obscure corners of human society may produce these "replications." The laboratory findings, then, would have prepared the way for an understanding of newly occurring or newly discovered social-psychological processes. It would be a poor science, indeed, that did not have some basis on which to make predictions about new situations before they occurred. For example, social psychology should be prepared to say something *now* useful in a world society in which all-out war does not occur.

The second type of answer to the question of generalizing from laboratory to field concerns the explanation for the causal sequence or law found in one or more studies. By explanation, the author means an account of the process postulated to mediate between cause and effect. The mediational process might involve a level of data different from that for the cause and the effect. This process can occur within the person in a form governed by individual-social laws or it can occur among a number of people (when a group law is involved). It should be obvious that the law would apply to cases for which there is independent ground for assuming that the mediational process could and does occur. For example, suppose the mediational process explaining the law is one that entails a person's seeking to perceive the world as consistently as possible. It would be expected that the laws could be gen-

eralized to any population for which there existed independent grounds for believing that there was motivation to achieve consistent perceptions.

The independent grounds can take a variety of forms: independent test results, knowledge of the cultural norms, knowledge of the early learning experience, or knowledge of their psychological structure. The original law would be strengthened, not weakened, if one found that the law did not hold in populations for which there was little or no reason to believe that the assumed mediating process was operative. The law would be strengthened in the sense that its explanation would have been independently supported. In short, the key to sound generalization from the laboratory to real life lies in the explanation of the laws based on laboratory findings.

This approach suggests also that, before predictions are made on the basis of laboratory findings, the researcher must ask himself not only what population of persons and of situations he believes he has sampled; but also ask himself what explanation he was using and what grounds exist for believing that the explanation holds in cases outside the laboratory. These grounds can come from levels of analysis and of scientific discourse different from the one at which the researcher works. For example, he might need to be concerned with economic factors that change child-rearing practices and cultural norms in such manner as to reduce the probability that the necessary mediating conditions will occur.

C. CONCLUSION

The general conclusion with regard to the three questions posed in this paper is that social psychology in its experimental form can make a unique contribution to the total scientific enterprise provided some thought is given to its role with regard to its neighbors.

REFERENCES

1. FESTINGER, L. Laboratory experiments. In L. Festinger & D. Katz (Eds.), *Research Methods in the Behavioral Sciences*. New York: Dryden, 1953.
2. FRENCH, J. R. P., JR. Experiments in the field settings. In L. Festinger & D. Katz (Eds.), *Research Methods in the Behavioral Sciences*. New York: Dryden, 1953.
3. HOVLAND, C. I. Reconciling conflicting results derived from experimental and survey studies of attitude change. *Amer. Psychologist*, 1959, 14, 8-17.
4. NAGEL, E. *The Structure of Science*. New York: Harcourt, 1961.

Department of Psychology
University of Washington
Seattle, Washington

ADOLESCENT RESPONSE TO RELIGIOUS APPEAL AS RELATED TO IPAT ANXIETY*

*Dallas Baptist Association and Houston Independent School District***

C. EWING COOLEY AND JERRY B. HUTTON

A. PROBLEM

Augustine has been described as having been a very anxious person. However, it has been noted that "One of the most remarkable effects of his conversion and subsequent career was a marked reduction of anxiety symptoms (2)." The area in religion referred to as "conversion," committal, or rededication is a fruitful but often inaccessible area for research.

The current study is concerned with two questions: (a) Are individuals with a high anxiety level more apt to respond to the emotional appeals of religious invitations? (b) Will individuals who respond to emotional appeals in a religious setting have a decrease in anxiety level?

B. METHOD

Seventy-two of 255 adolescents attending a Southern Baptist youth camp¹ responded to a religious appeal by moving to the front of the tabernacle following a series of preaching services and during a special invitation period. The campers were from Houston and surrounding suburbs. They represented practically every section of town and ranged from lower to higher socioeconomic status. The 72 adolescents who made a response were responding to a call for conversion, rededication, or special-service. Most of the decisions were for rededication. Both the total number of campers and the number making decisions closely approximated 43 per cent males and 57 per cent females. Table 1 compares all campers with those who made response.

On the first day of camp, during an afternoon rest period, the IPAT Anxiety Scale Questionnaire (Self-Analysis Form) was given to all student campers. Those who made responses to appeals throughout the week were given the test again on the last day of camp, three days later. Comparisons were made between the test scores before and after the response.

It was not practical to include a control group in this study as the purpose

* Received in the Editorial Office, Provincetown, Massachusetts, on June 1, 1964. Copyright, 1965, by The Journal Press.

** Now at Mental Evaluation Clinic, Houston, Texas.

¹ The authors are indebted to Cortland Savage, Camp Director and Youth Director at Park Place Baptist Church, Houston, Texas, for allowing and securing associational support for this study.

TABLE 1
AGE AND SEX COMPARISONS OF THE TOTAL NUMBER OF CAMPERS
WITH THOSE WHO MADE RESPONSES

Population and sex	13*	14	15	Age 16	17	18	19	Total
All campers								
Male	10	25	24	31	16	1	3	110
Female	10	38	32	36	24	4	1	145
Total	20	63	56	67	40	5	4	255
Campers who responded								
Male	3	8	9	6	4	0	1	31
Female	2	14	7	6	9	3	0	41
Total	5	22	16	12	13	3	1	72

* Those who were 13 years of age were within three months of being 14.

of the Harris County Baptist Association was to invest in testing primarily as an aid in screening and group work. Funds were not available to obtain pretests and posttests from a comparable group of adolescents who did not attend camp or from those in the camp who made no response.

C. RESULTS

Of the 255 campers, 55 had a sten rating of 8 or above indicating a high anxiety level. This involved 22 per cent of the total number of campers. Fourteen of the 72 who made responses had a sten rating of 8 or above, representing 19 per cent with high anxiety. Thus, it appears that those with high anxiety ratings did not tend to respond to religious appeals to a greater extent than did those with low anxiety ratings.

The total number of campers had a sten mean of 6, while those who made responses had a sten mean of 6.2. Again, it appears that those who made responses did not tend to have a significantly higher level of anxiety than the total number of campers.

Based on the frequencies of males and females who increased, decreased, or remained the same in raw-score anxiety level following response to a religious appeal, chi squares were found to be 24.9 for males ($p < .001$), 5.4 for females ($p < .02$), and 22.8 ($p < .001$) for all responders.

The test-retest reliability for the three-day interval was .78. The pretest standard deviation was 9.7 (mean = 36) and the posttest standard deviation was 10.8 (mean = 31). A critical ratio of 6.1 shows that the mean on the posttest is significantly lower (indicating lessened anxiety) than the pretest mean.

D. DISCUSSION

One can conclude that adolescents who respond to invitations for conversion, rededication, or committal to special service are not necessarily individuals with high levels of anxiety. It follows that the emotional appeal of a

religious service does not tend to cause some individuals with high anxiety levels to respond to a greater degree than those with other anxiety ratings. In other words, anxiety does not seem to be related to a response to a religious appeal.

It should be noted that although it appears that anxiety was significantly lowered among those who responded, it is not known if the remaining campers demonstrated changes in the level of anxiety. Perhaps, due to the total experience of the camp, all of the campers experienced a lessening of anxiety level. A possible restudy is suggested that would give a second test to those who did not respond, and a first and second test to those who did not attend camp. A comparison could then be made between the three groups to see if the total camp experience or response to a religious appeal might be influencing factors.

By observing the changes in the various areas of anxiety as subdivided by Cattell and Scheier (1), it appears that the lowering of anxiety occurs to the greatest extent in the areas of suspiciousness, paranoid insecurity, and guilt proneness. The decrease does not seem to occur as significantly in areas related to personality structure. Anxiety related to ego strength is not lowered significantly. Although the part scores are too brief to be of significance, it might be said that the reality of the situation has not changed. The individual still has to face the same problems in life.

E. SUMMARY

Seventy-two adolescents, among 255 attending a summer youth camp in a religious setting, responded to group invitations for rededication, conversion, or committal to special service. The total number of campers were given the IPAT Anxiety Scale Questionnaire on the first day of camp, and the 72 responders received it the second time on the last day of camp. The comparison of pretests and posttests suggests that levels of anxiety are not significantly related to responses to religious appeals. Furthermore, it appears that those who responded to the appeals experienced a lessening of anxiety to a significant degree.

REFERENCES

1. CATTELL, R. B., & SCHEIER, I. H. Handbook for the IPAT Anxiety Questionnaire. Champaign, Ill.: Institute for Personality and Ability Testing, 1963.
2. OUTLER, A. C. Anxiety and grace: An Augustinian perspective. In S. Hiltner & K. Menninger (Eds.), *Constructive Aspects of Anxiety*. New York: Abingdon, 1963. P. 94.

Mental Evaluation Clinic
6621 Fannin
Houston, Texas 77025

FAMILIARITY, DOGMATISM, AND REPORTED STUDENT ATTITUDES TOWARD THE DISABLED*¹

Rehabilitation-Education Center, University of Illinois

JACK K. GENSKOW AND FRANK D. MAGLIONE

A. INTRODUCTION

Phenomena related to attitudes of persons toward the disabled should, it would seem, be understandable and capable of description in terms of the psychology of attitudes in general. When given two milieus known to differ in a particular characteristic suspected as capable of affecting attitude expression, and when introducing specific conditions also suspected as capable of affecting such expression, data obtained by means of devices purporting to measure expressed attitudes should "behave" in generally anticipative fashions. More specifically, one might expect differences in expressed attitudes toward the disabled between a campus population having considerable familiarity with the disabled and a second having little familiarity. Further, differences might be expected within each population if the measuring devices were administered by a person in a wheelchair and by a person not in a wheelchair. It was believed that attitudinal data obtained under these varying conditions could be described in terms of more inclusive behavior characterizations.

Equilibrium theory (4, 7) suggests the existence of a tendency toward consistency between cognition, perception, affect, and behavior. Consistency among these elements would be expected with regard to attitudes toward the disabled; for example, an increase in positive or negative affect in the presence of a disabled individual would tend to be reflected in the cognitive behavior of a person completing an unsigned attitude questionnaire dealing with the disabled. Rosenberg *et al.*'s concept of "threshold for inconsistency" (11) suggests that in some situations a person's attitude (or reported attitude) would be more consistent with its cognitive, affective, and behavioral elements than in other situations. Thus, with a disabled individual present, a subject's anonymously reported attitude might be more consistent with his perception, cognition, and affect than in a less reality-bound situation when no disabled individual is present. Heider's "balance theory of sentiments" suggests that

* Received in the Editorial Office, Provincetown, Massachusetts, on June 12, 1964. Copyright, 1965, by The Journal Press.

¹ This research was performed while both authors were doctoral candidates in Education. The authors wish to acknowledge the assistance of Professors T. E. Newland and C. H. Patterson, and also Mrs. Patricia Griffin for her work on the early stages of this study.

unfamiliarity, as well as perceived lack of similarity, between the perceived and the perceiver, tends to yield a negative attitude. Hebb (6) stated that an object (e.g., a disabled person) which is partly familiar—i.e., which presents stimuli that arouse “habitual processes of perception” but that also is different in enough respects to elicit “incompatible processes”—tends to produce fear. Thus attitudes toward the disabled could be posited to depend in part upon the perceiver’s previous contacts or familiarity with the disabled, as well as upon perception of similarity—dissimilarity.²

Rokeach (10) was concerned with the structure or organization of belief systems (e.g., tolerance of others) rather than with the specific content of the belief system. He differentiated between individuals who are relatively open (low dogmatic or low authoritarian) and individuals who are relatively closed (high dogmatic, high authoritarian) with respect to belief-disbelief systems about reality. Rokeach suggested that the highly dogmatic person tends to see his environment as threatening and (as a result) is intolerant of others, including (the present authors assume) the physically disabled.

In terms of such theoretical orientations, the following hypotheses were examined.

1. Two groups believed to differ in familiarity with the disabled would be expected to respond differently to a questionnaire containing items purporting to measure attitudes toward the disabled. Thus the group more familiar with the disabled would be expected to report a more positive attitude toward the disabled than would the group which was less familiar with them.

2. Further differentiations between the groups could be made between questionnaire administrations by an ostensibly disabled person and an ostensibly nondisabled person; i.e., with the “disabled” administrator present, the group more familiar with the disabled would react more positively toward the disabled; and the less familiar group would react more negatively toward them than they would when no “disabled” person was present.

3. In view of the possible dynamic relationship between dogmatism and the expressed attitudes, it was hypothesized that items purporting to measure degree of dogmatism would show a relationship to attitudes—with high dogmatism being associated with negative attitudes expressed toward the disabled.

B. METHOD

The subjects were 111 students (approximately 60 of whom were female) in two educational-psychology classes at each of two state universities: Indiana and Illinois. The 57 students at the University of Illinois were juniors and

² See Wright (13) for a further discussion of Hebb's and Heider's theories applied to attitudes toward the disabled.

seniors who had the same instructor. One Indiana University class was composed of upperclassmen; the other class, with a different instructor, consisted of underclassmen. Because the University of Illinois has had an extensive program for physically disabled students for a number of years and because Indiana University does not have such a program, the Illinois subjects are designated as the familiar group (F); the Indiana subjects, as the unfamiliar group (U).

Two questionnaires were combined: the Attitude Toward Disabled Persons Scale (ATDP), developed by the Human Resources Foundation (14), and Rokeach's Dogmatism Scale (DS), Form E (10). Two combined forms of 40 items each were used. The ATDP-DS-1 was composed of all 20 ATDP items randomly combined with the odd-numbered items of the 40-item DS; the ATDP-DS-2 randomly combined the same 20 ATDP items with the remaining 20 DS items. The same tape-recorded directions were used for all questionnaire administrations.

Four mixed-sex classes were used, two at each university. Each class was first administered the ATDP-DS-1; and, after an interval of one or two weeks, the ATDP-DS-2. Each of the four able-bodied administrators gave the questionnaire twice to different classes: once as an able-bodied individual; once while posing as a "disabled" individual in a wheelchair. Table 1 shows the procedural design.

After each instructor announced to his class that cooperation in some research was desired, he turned the class over to the administrator, who proceeded to distribute the questionnaires from the front of the room. Then he turned on a tape recorder and played directions that were used for every administration. When all questionnaires were complete, the administrator col-

TABLE 1
PROCEDURAL DESIGN

Administration	Group U*		Group F**	
	Class A	Class B	Class C	Class D
First (ATDP-DS-1)	A-1	B-1	C-1	D-1
	Administrator One Able-bodied	Administrator One Wheelchair	Administrator Two Able-bodied	Administrator Two Wheelchair
Second (ATDP-DS-2)	Administrator Four Wheelchair A-2	Administrator Four Able-bodied B-2	Administrator Three Wheelchair C-2	Administrator Three Able-bodied D-2

* Two weeks elapsed between the first and second test administrations.

** One week elapsed between the first and second test administrations.

lected them and left the room. The entire procedure took less than the total class period of 50 minutes.

The tape-recorded directions were as follows:

We are using this tape recorder to standardize the administration of this questionnaire. In other words, each group to which this questionnaire is given will hear these same directions given the same way over the tape recorder. This questionnaire is the first of a two-part research design, the results of which will be used for research purposes only. The second part of the scale will be given to you at a later date. The questionnaire is designed to measure your attitudes and opinions toward several different factors.

Please do not put your name or other identifying marks on the questionnaire; this study is interested in the results of the class as a whole rather than individual results. If any of you know, or have seen the person administering the test before, please raise your hand [pause]. It is believed that the directions are self-explanatory. Please begin now and proceed at your own pace.

To analyze the data the *t* test for independent samples was used with both the ATDP and the DS on all comparisons between groups U and F (i.e., total group comparisons, intergroup wheelchair administration comparisons and intergroup able-bodied-administration comparisons), and also on the ATDP test-retest comparisons within both groups under the two administration conditions: i.e., wheelchaired *vs.* nonwheelchaired administrators. The more powerful and appropriate *t* test for matched samples could not be used in the test-retest situations because the individual questionnaires were anonymous and unidentifiable for individual comparisons between administrations. Anonymity also accounted for the *N*s being different for each administration. Relationships between the ATDP and DS (split-half) forms for each administration were determined by means of Pearson *r*.

C. RESULTS

Using *t* tests for independent samples, it was found that the group of college students more familiar with the physically disabled (F) reported a more positive attitude toward the disabled than did the group less familiar with them (U), the difference being significant beyond the .0005 level [see Table 2]. The two groups were seen to differ significantly also on the dogmatism variable, Group F reporting a less dogmatic approach ($p < .01$). [See Tables 2 and 3.]

Able-bodied-administration comparisons between Groups U and F on the ATDP differ at the .15 level of significance; whereas between-group wheel-

TABLE 2
ATDP AND DS TOTAL-RESPONSE COMPARISONS

Instrument	<i>N</i> *	Group U Mean	<i>SD</i>	<i>N</i> *	Group F Mean	<i>SD</i>	Difference	<i>t</i>	<i>p</i> **
ATDP	109	80.01	11.84	113	86.39	12.15	6.38	3.643	.0005
DS	109	75.48	12.78	113	70.12	11.14	5.36	2.807	.01

* *N* is based on two questionnaire administrations to each group.

** ATDP, one-tailed *t* test; DS, two-tailed *t* test.

TABLE 3
INTERGROUP COMPARISONS OF WHEELCHAIR AND ABLE-BODIED TEST ADMINISTRATIONS

Administration		N**	ATDP		DS	
Group U	Group F		<i>t</i>	<i>p</i> *	<i>t</i>	<i>p</i> *
AB	AB	107	1.196	.15	2.014	.05
WC	WC	115	3.508	.0005	2.410	.02
WC/AB	WC/AB	222	3.643	.0005	2.807	.01

* ATDP, one-tailed *t* test; DS, two-tailed *t* test.

** *N* is based on two questionnaire administrations to each group.

chair administrations differ at the .0005 level [see Table 3]. These data are in accord with the second hypothesis, which is that the wheelchair-administration means for the two groups would differ in opposite directions from their respective able-bodied-administration means. However, Table 4 indicates that—whereas Group U's wheelchair-administration mean was lower (less favorable) and Group F's was higher (more favorable) than their respective able-bodied-administration means—as predicted these differences are significant only at the .15 level. If the more appropriate and more powerful *t* test for paired samples had been used, the results might have reached a more significant level. The hypothesis that wheelchaired and able-bodied administrators effect a change in the reported attitude must be rejected.

No product-moment correlation coefficients approaching significance were found between measures of dogmatism and attitudes toward the disabled [see Table 5].

D. DISCUSSION

Heider's view that both the unfamiliar and the dissimilar yield negative reactions appears to be applicable to attitudes toward the disabled. For Group U, which has had little or no previous contact with the disabled, the disabled were probably both unfamiliar and dissimilar and thus perceived in a negative light. Group F would be more familiar with the disabled, and the perceived similarity—even if only in terms of being fellow students on the same campus—would also increase between the able bodied and the disabled. This increased familiarity and similarity would partially account for the more favorable attitude expressed by Group F. Another factor in Group F's more positive reported attitude is the fact that the nature of contacts with the disabled students at the University of Illinois is probably more positive than negative because the disabled students are rigorously screened prior to admittance to the University, and thus constitute a positively biased sample of the disabled population.

TABLE 4
INTRAGROUP COMPARISON OF WHEELCHAIR WITH ABLE-BODIED TEST ADMINISTRATIONS OF THE ATDP

Group	Wheelchair				Able-bodied				Difference	<i>t</i>	<i>p</i>
	Classes	<i>N</i>	Mean	<i>SD</i>	Classes	<i>N</i>	Mean	<i>SD</i>			
U	B1 + A2	57	78.69	12.69	A1 + B2	52	81.83	14.38	3.14	1.208	.15
F	D1 + C2	58	87.68	13.20	C1 + D2	55	85.02	13.39	2.66	1.087	.15

TABLE 5
MEANS, STANDARD DEVIATIONS, AND INTERSCALE CORRELATION COEFFICIENTS
FOR EACH ADMINISTRATION OF ATDP-DS

Administration	Class	N	ATDP*		DS**		r***
			Mean	SD	Mean	SD	
Group U							
Able-bodied	A-1	27	83.14	14.82	79.66	12.89	— .049
Wheelchair	A-2	26	79.15	12.84	76.03	11.41	— .028
Wheelchair	B-1	31	78.25	12.89	74.67	10.87	— .027
Able-bodied	B-2	25	80.40	13.86	71.52	14.29	— .047
Group F							
Able-bodied	C-1	25	84.16	12.84	74.00	9.68	— .005
Wheelchair	C-2	26	87.03	10.30	68.57	11.41	— .028
Wheelchair	D-1	32	88.84	9.96	70.43	10.26	— .037
Able-bodied	D-2	30	85.66	14.26	68.00	12.55	— .015

* A high score indicates favorable attitude toward the disabled.

** High scores indicate high dogmatic. These are split-half forms of the DS.

*** Pearson product-moment correlation coefficients between the ATDP and DS.

Equilibrium theory would account for the difference in reported attitudes between Groups U and F in terms of past inconsistencies in the attitude structures of Group F, these inconsistencies being brought about by favorable experience with the disabled. According to Festinger, if the dissonance created by these inconsistencies became greater than the resistance to change, then the least resistant elements of cognition would be changed; thus reducing the dissonance. Thus, whereas the precollege attitudes of Group U and Group F may well have been similar, inconsistencies in Group F were brought about by contact with disabled students which these results suggest may have resulted in a positive change in attitude more consistent with experience with the disabled.

These results are in agreement with other relevant studies (14), which used the ATDP and which also indicated that reported attitudes toward the disabled were more favorable for groups having had previous contact with the disabled. More research is needed before the processes underlying this relationship between familiarity and acceptance can be more fully explained. It should be noted that when discussing the difference in attitudes expressed by Groups U and F, it is not necessary to correct for the difference in dogmatism between the two groups because attitudes and dogmatism were seen to be essentially not related [see Table 5].

Although the hypothesis that wheelchaired and able-bodied administrators would yield attitude responses in opposite directions for Groups U and F is tentatively rejected, the results were in the expected direction. Furthermore, the hypothesis was rejected using a *t* test for independent samples in a test-

retest situation; the more powerful t test for matched samples possibly would have yielded a larger t and a stronger confidence level (by taking into account the correlation between test and retest). Also, whereas the interschool difference in reported attitudes was barely significant when the able-bodied administrations of the ATDP were compared, when the interschool wheelchair administrations were compared the difference was found to be highly significant. Thus it appears that the presence of a "disabled" individual accentuated the difference in reported attitudes between Groups U and F to some degree, which was predicted. Granting that further research is needed on this hypothesis, the present findings merit further examination.

As stated earlier, both Heider and Hebb suggest that the unfamiliar and the partly familiar, which the authors presume the disabled to be, are "fear inducing." Thus, Group U, which has had little or no contact with the disabled, would be expected to experience a rise in affect in the presence of a physically disabled individual. Subjective confirmation of this increase in affect was supplied by a questionnaire administrator at Indiana University when he compared the rapport established with two classes when he was in a wheelchair and when he was able-bodied. He characterized the wheelchair class as being "embarrassed, unwilling to look me in the eye." In terms of equilibrium theory, when the disabled individual was present with Group U and a resultant rise in negative affect occurred, the reported attitude under these changed conditions would reflect this change in affect; i.e., the behavioral and cognitive attitude components as reported in the questionnaire would become more consistent with the perceptual and affective components. This would suggest that the tendency (at the .85 confidence level) for Group U to express a more negative attitude toward the disabled when the disabled person was actually present than when he was not present was a function of rise in negative affect.

The wheelchair administrators representing ambiguous (or perhaps even stereotyped?) disabled persons would tend to be perceived largely as a function of the perceiver's own expectations based on past experiences. For Group F the disabled administrator would serve as a stimulus to "remind" them of prior contacts with disabled students on campus.

Rosenberg *et al.* (11, p. 55) suggest that when "affective-cognitive inconsistency is established in a context which forces the person to think closely about it [as when one is completing an attitude-toward-the-disabled questionnaire in the presence of a disabled individual] then the intolerance for inconsistency is maximized." This theory posits that the cognitive reporting of an attitude would reflect the nature of the heightened affect. Group U would

report a more-negative attitude; Group F, a more positive attitude. Rosenberg's response threshold to inconsistency states that one can tolerate inconsistencies in his attitude structure in some situations better than in others. For both groups the presence of a wheelchair administrator might have raised any inconsistency in attitude above the response threshold, thereby producing a changed response. Conversely, when no disabled person was present, the situation was less reality bound, and any inconsistencies in attitude were below the response threshold. This suggests that the reported attitude when the disabled administrator was present was actually more consistent with the groups' "actual" attitudes toward the disabled than was the attitude reported when no disabled individual was present.

The extremely low correlation between dogmatism and reported attitudes toward the disabled may be a function of the interaction of several different variables. In his studies, Rokeach tended to use only the top and bottom groups of scores (i.e., the extreme high and low dogmatics), whereas this study used all scores. Second, the use of a split-half version of the DS probably affected the general reliability of the instrument and thus may have contributed to a lesser correlation. Third, Weitman (12) makes a distinction between various kinds of authoritarians and points out that some individuals have indifferent or ambivalent attitudes toward those items in the scale designed to measure dogmatism. If the sample chosen for this study was similarly nonego involved, it would contribute to the rejection of the dogmatism-attitude relationship; such lack of ego involvement, however, would not necessarily carry over to the items measuring attitudes toward the disabled because of the apparent perception by the subjects of the test situation as one dealing with disabled persons.

Bruner and Tajfel's discussion (3) of Arnhoff's failure (1) to replicate an earlier study dealing with the tendency of ambivalent ethnocentric personalities to overgeneralize suggested that the dogmatic individual may either overgeneralize or undergeneralize, depending upon situational factors as perceived at the time. This inconsistency in the results of previous studies therefore points up the possible existence of numerous unanalyzed variables, which in the present study may account for the lack of correlation between dogmatism and attitudes toward the disabled. The results of this study do agree, however, with the correlation of .13 between the ATDP and the DS, based on a sample of 103 college students, reported by one of the authors of the ATDP.³

The following possible influencing factors should be recognized: (a) In

³ Personal Communication from J. R. Block, July, 1962.

class discussions following each final administration, several members of one of the Group F classes pointed out that the wheelchair examiner was probably not disabled since his legs were not atrophied and his shoes showed signs of normal wear. Several other members of this same class observed that they detected "slowness" in wheelchair maneuvering, even though the administrators had been trained in the use of wheelchairs by rehabilitation personnel. Such comments reflect familiarity with fellow students who are members of the Illinois Student Rehabilitation Program. (b) One class at the University of Illinois included a student disabled with moderately severe cerebral palsy who did not use a wheelchair. This student was not present during the questionnaire administration. (c) Comments were made by members of the two Group F classes to the effect that they perceived both questionnaires (ATDP-DS-1 and ATDP-DS-2) as being identical and as obviously concerned with disabled persons. (d) Since the ATDP and the DS were randomly combined, a potential interactive effect might be assumed. Thus the taking of the ATDP might have affected scores on the DS and vice versa. Since the combined questionnaires were perceived by some subjects as "having something to do with the disabled" and since the same ATDP items were repeated on both forms of the combined scales (whereas split-half DS items were used), it might be presumed that the ATDP had a greater effect on the DS than vice versa. (e) It has been suggested that the DS is subject to an acquiescent-response set (8, 9). (f) Except for the "disabled"-nondisabled variable, the personality and general stimulus value of the administrators were purposely kept neutral. It is suggested, therefore, that the wheelchair administrators were viewed as "stereotypes": i.e., a function of the perceiver's own viewpoint toward disabled persons (2). (g) Although all University of Illinois students see physically disabled students nearly every day, the amount of contact varies for each student. This difference in contacts was not measured in this study. (h) Since contact with disabled students at Illinois is presumed to be mostly positive, a learning-theory approach to these results could possibly be used to account for the more-positive reported attitude at Illinois. (i) Generalizations beyond this study presumably should be limited to comparable college samples.

E. SUMMARY

Student attitudes toward the disabled and dogmatism were investigated in two midwestern universities: one having had an extensive program for physically disabled students for many years; the other having had no such program. On the basis of "equilibrium theory," Hebb's neurophysiological theory, and Rokeach's work on dogmatism three hypotheses were tested. Two classes at

each university were on two occasions administered different combined forms of the Attitude Toward Disabled Persons Scale (14) and the Dogmatism Scale, Form E (10), once by a person in a wheelchair and once by a person not in a wheelchair. Tape-recorded directions were used. Between-school and within-school results were considered in terms of correlations between the scales and *t* tests between the administrations.

The first hypothesis was supported: Students familiar with disabled persons reported a significantly more favorable attitude toward the disabled than did students unfamiliar with them. The second hypothesis was tentatively rejected, although the results were in the expected direction; i.e., when a wheelchair administrator was present, Group F reported a more positive attitude toward the disabled. Group U reported a more negative attitude than when no "disabled" administrator was present. Differences between the two groups were considerably more significant for wheelchair administrations than for able-bodied administrations. Although no significant correlations were found between dogmatism and attitudes toward the disabled, Group F demonstrated a significantly less dogmatic, more open-minded attitude structure.

These results were interpreted as being in general agreement with equilibrium theorists as well as with Hebb. The results regarding dogmatism, while inconclusive, were not in agreement with Rokeach's theory.

REFERENCES

1. ARNHOFF, F. N. Ethnocentrism and stimulus generalization: A replication and further study. *J. Abn. & Soc. Psychol.*, 1957, **55**, 393-394.
2. BRUNER, J. S., & POSTMAN, L. An approach to social perception. In W. Dennis (Ed.), *Current Trends in Social Psychology*. Pittsburgh: Univ. Pittsburgh, 1948. Pp. 71-118.
3. BRUNER, J. S., & TAJFEL, H. Cognitive risk and environmental change. *J. Abn. & Soc. Psychol.*, 1961, **62**, 231-241.
4. FESTINGER, L. A Theory of Cognitive Dissonance. Stanford, Calif.: Stanford Univ. Press, 1957.
5. GENSKOW, J. Dogmatism and Reported Student Attitudes Toward the Disabled. Unpublished Master's thesis, University of Illinois, Urbana, 1962.
6. HEBB, D. O. On the nature of fear. *Psychol. Rev.*, 1946, **53**, 259-276.
7. HEIDER, F. *The Psychology of Interpersonal Relations*. New York: Wiley, 1958.
8. LICHTENSTEIN, E., QUINN, R. P., & HOVER, C. L. Dogmatism and acquiescent response set. *J. Abn. & Soc. Psychol.*, 1961, **63**, 636-639.
9. PEABODY, D. Attitude content and agreement set in scales of authoritarianism, dogmatism, anti-semitism and economic conservatism. *J. Abn. & Soc. Psychol.*, 1961, **63**, 1-11.
10. ROKEACH, M. *The Open and Closed Mind*. New York: Basic Bks., 1960.
11. ROSENBERG, M. J., HOVLAND, C. I., MCGUIRE, W. J., ABELSON, R. P., & BREHM, J. W. *Attitude Organization and Change: An Analysis of Consistency Among Attitude Components*. New Haven: Yale Univ. Press, 1960.

12. WEITMAN, M. More than one kind of authoritarianism. *J. Personal.*, 1962, **30**, 193-208.
13. WRIGHT, B. A. *Physical Disability—A Psychological Approach*. New York: Harper, 1960.
14. YUKER, H. E., BLOCK, J. R., & CAMPBELL, W. J. A Scale to Measure Attitudes Toward Disabled Persons (Human Resources Study No. 5), Human Resources Foundation, Albertson, N. Y., 1960.

Rehabilitation-Education Center
University of Illinois
MRH Post Office Box 517
Champaign, Illinois 61822

DEMAND CHARACTERISTICS VS. PHYSIOLOGICAL NEEDS:
A METHODOLOGICAL CRITIQUE OF THE "NEW LOOK"
SCHOOL OF PERCEPTION*¹

Department of Psychology, University of California, Berkeley

ARLENE C. BRANDON AND EDWARD E. SAMPSON

A. INTRODUCTION

The purpose of this study is to demonstrate the importance of demand characteristics as a confounding factor in psychological experiments, particularly in those experiments that deal with the "new look" in perception.

It is not a newly discovered fact that in many experiments the subjects respond not only to the manipulation of the experimental variable but also to the implicit expectations of the experimenter (8). In general, the conception of *S* as a helpful, cue-seeking individual has been ignored by most *Es* who claimed immunity to this bugaboo through the inclusion of the omnipotent "control group." Recently, however, Orne (9, 10) has discussed the far-reaching effects of what he calls demand characteristics and has emphasized that such characteristics cannot be adequately treated by the typical control group. The effects of demand characteristics have been demonstrated in sensory-deprivation experiments (6, 11) and in an experiment on cognitive consistency (15). These studies have demonstrated that *Ss* can pick up cues from the experimental situation and, on that basis alone, respond in a manner consistent with the experimental hypotheses being tested.

Some experiments are more susceptible to demand-characteristic responses than others. The Atkinson and McClelland study in perception (3) seems to be representative of those that are easily affected. In that study, *Ss* were restrained from eating for as long as 16 hours. They were then brought into the experimental room where they were asked to respond to slides (blanks) projected on a screen in terms of various kinds of questions (some of which were of the multiple-choice type) that included one food item as a possible choice. The results indicated that under periods of long deprivation, *Ss* reported more food and food-related items than under periods of short deprivation. Furthermore, *Ss* reported more food items when the questions were

* Received in the Editorial Office, Provincetown, Massachusetts, on June 15, 1964. Copyright, 1965, by The Journal Press.

¹ A brief version of this paper was presented to the Western Psychological Association, April 1964.

more tightly organized; i.e., the multiple-choice questions pulled more food responses than the projective questions. The authors accounted for this phenomenon in terms of the strength of the "hint" implicit in the question. Following Orne's suggestion, we hypothesize that the "hint" existed not only in the form of the questions but throughout the entire experimental design. The responses were not only a result of *Ss*' physiological condition but also of their motivation to help *E* in a situation involving clear expectations. Because of this possible confounding the Atkinson and McClelland results remain equivocal. Further research in this area (7) has improved the research design; however the original design is an excellent model for our demonstration.

B. METHOD

The main methodological difficulty in an experiment on demand characteristics is to change the demand characteristics without changing anything else. To do this we used the well-established placebo effect (13). In the past, placebos have been used clinically (2, 12) and in hypnosis studies (4, 5). In both cases the results produced were convincing because great care was taken to establish the credibility of the situation as one in which a drug conceivably could be given.

1. Subjects

Ss were 39 male and female undergraduates at the University of California enrolled in their first psychology course. Participation in an experiment was a requirement of the course.

2. Procedure

a. Experimental condition. The procedure started at the time *S* signed up for the experiment, at least one day prior to the testing. Near the sign-up sheet was a prominent notice requesting the prospective *S* to eat a large breakfast on the morning of the experiment—in fact, to eat a great deal more than usual. (Apparently a self-selection procedure operated because no coeds on diets would sign up for the experiment.) When *S* arrived at 10 a.m. the next morning he entered a room filled with medical-looking odds and ends; a poster describing the intricacies of the Krebb's Cycle, books with titles such as *The Hypothalamus* and *The Living Brain*, a table set up as a cot, and an *E* dressed in a white coat with a stethoscope dangling from her pocket. After the preliminary greeting, *E* gave *S* a description of the experiment.²

² Copies of the complete description of the experiment given to *Ss*, the three release forms, the medical questionnaire, the placebo questionnaire, the instructions for the slides, and the questions used for the slides may be obtained from the senior author.

The description indicated to *S* that he was about to receive a new and unusual kind of pill that would put his body in a state of hunger, even though he might not actually *feel* hungry. This pill was said to be as unusual as the psychomimetic drugs in that it could cause drastic effects in a short period of time. The explanation was as follows:

What we have for you today is not quite as dramatic as the hallucinogens but just as unusual. As you know, your hunger center is not in your stomach but in your brain, in the ventromedial nuclei of the hypothalamus. When did you eat and what did you eat? Good. That means you now have a high blood-sugar level. That is a controlled part of this experiment. Now if I were to give you a pill that would reduce your blood-sugar level suddenly and drastically, the ventromedial nuclei would start to send hunger messages to your digestive system, and eventually you would feel hungry, your stomach would start to contract and secrete gastric juices. I have just such a pill. I am particularly interested in the period of time between the activation of the ventromedial nuclei and the actual activity of the digestive tract: that is, the period when your body is in a state of hunger but you do not yet feel hungry. That is much like the neurons firing in your brain. They are active, but you are unaware of them; so the experiment consists of showing you some slides, giving you the pill, and then showing you another set of slides.

At this time *S* was given the option of leaving the experiment right then while still receiving his credit as a participant, for *E* would use only an *S* who voluntarily stayed and who would sign some release forms.

The *Ss* who agreed to stay (and all but one did) read and signed three forms. The first released the University of California and all affiliated personnel from any responsibility for *S*'s welfare during the experiment; the second released *S*'s medical records to *E*. The third release form stated that the drug was harmless if *S* had eaten within the last hour and a half and required *S* to swear to the fact that he had eaten within that time. Next, *S* was asked for a brief medical history that consisted of 26 questions of the following type: "Have you ever had or do you now have rheumatic fever, TB, etc." "When was the last time you were in a hospital and why?" "When was the last time you saw a doctor and why?" Shortly after the foregoing questionnaire, *E* took *S*'s temperature and pulse. A placebo questionnaire, based on the one used by Abramson (1), was then given with the following introduction:

I am interested in any slight changes that may occur within you due to the pill; therefore, I will ask you these questions now and then again after the pill. You may not be feeling exactly normal right now because you are in a strange situation with an unknown pill coming that you are not very sure of. (Most *Ss* smiled and looked embarrassed at this point.) I want you to answer these questions using "yes," "no," or "maybe" as

answers. Compare yourself now to your usual condition—let us say, as you were an hour ago. The “maybe” category is very important because it is unlikely that there will be any great dramatic changes but only subtle ones; thus, please try to be introspective and sensitive about the changes.

b. Control-drug condition. Ss in the control-drug condition were then shown the first set of slides. The directions for the slides were only a slight modification of those used by Atkinson and McClelland. These directions state that the purpose of the experiment is to measure visual acuity under different physiological conditions. The instructions also emphasize over and over again the difficulty of the task and the need for what seems to be pure guessing. We added the following statements to the instructions:

There are two kinds of slides. One is a white slide with a thin line drawn on it. The other is a blotchy black-and-white slide with a thin line drawn on it. In most cases you can tell when you have a white slide and when you have a blotchy slide—but not all the time. That's how difficult they are. The best thing to do is to read the question, look at the slide and write without too much thinking in between. It is not legitimate to write “white slide, blotchy slide, white light, etc.” You must name an object.

Next, *E* showed *S* how the slide projector and attachments worked and three sample questions: a multiple-choice question, a projective question, and a loosely structured question. After making sure that *S* had no questions about the clerical procedure the first set of slides was projected on the screen. When all 30 slides had been presented, the placebo was administered.

S was given two yellow capsules from a bottle marked “For Experimental Use Only” and a glass of water. A stopwatch was started and *S* was asked to lie down on the table for 10 minutes, during which time he was to think only about himself. *E* called this “socially approved thinking about yourself.” Again, *E* emphasized that *S* should not expect any dramatic changes but should be sensitive and introspective enough to report any of the subtle changes. After five minutes, *E* took *S*'s pulse and said, “Uh-huh.” *E* told *S* when six minutes and nine minutes had passed. After 10 minutes *S* returned to his original seat and answered the same set of questions, this time with the instructions to compare his present feelings with his feelings before he took the pill. At the conclusion of the questioning *S* was asked to describe his feelings in his own words or to mention anything not covered by the questions. *S* was then shown the second set of slides. An interview followed and was aimed at discovering *S*'s knowledge of the hypothesis seemingly being tested, the effect of the pill, and belief in the pill. [See Table 1 for a detailed description of the steps in all four conditions.]

TABLE 1
DESIGN OF THE EXPERIMENT
(Numbers indicate the sequence of the procedure for each condition)

Condition	Directions at time of sign-up to eat large breakfast	Signing of three release forms	First placebo question- naire	First set of slides	Pill	Rest	Second placebo question- naire	Second set of slides	Second pill	Second rest
Control-drug	1	2	3	4	5	6	7	8		
Drug-control	1	2	3	7	4	5	6	10	8	9
Control no-pill	1	2	3	4		5	6	7		
Pure control			1	2		3	4	5		

c. Drug-control condition. Subjects in the drug-control condition went through the same introductory phase but were given the pill immediately after the first questionnaire, rested for 10 minutes and then were shown the first set of slides. After that, they were given the following information:

It would take 20 minutes for these pills to wear off, which is necessary before I can show you the second set of slides. But since you would probably like to be finished as soon as possible, I will give you this pill that will counteract the first and return you to your former physiological state.

S was given one pink capsule and asked to lie down for another 10 minutes; then he was shown the second set of slides and interviewed.

The true nature of the experiment was revealed to all *Ss* after the interview. All expressed astonishment at the elaborate measures used in the study. None seemed angry or upset at being the target of such procedures.

d. Pure-control group. The first control group was a pure-control group. These *Ss* did not undergo any of the elaborate procedures mentioned earlier. This group simply signed up for a one-hour experiment. There was no notice of food; so they came to the experiment unaware that it was concerned with food responses. *S* took the placebo questionnaire, the preface being that he may be feeling unusual because he was in an experimental situation. *S* was shown a set of slides, asked to lie down for 10 minutes, given the questionnaire a second time, and then shown the second set of slides.

This procedure gives a base rate for three measures: the number of food responses given to each set of slides; the difference in the number of food responses given at the first and second viewing; and the number of placebo changes that can be ascribed only to changes over time, a change common for all groups.

e. Control no-pill group. We hypothesized that *S* would perceive more food items in the blank slides, according to the demand characteristics of the experiment; that is, over two sets of slides the food responses would increase if the expectations were in that direction and would remain the same if expectations were at a minimum. Would food responses decline or remain level when this was the requirement of the demand characteristics? To answer this question, a second control group was tested. *Ss* in the control no-pill condition thought they were in the experimental condition and, in fact, were treated exactly as were *Ss* in the control-drug condition up until given the pill. At that time, instead of giving *S* the pill, *S* was told that he was in a control condition and would not receive the drug; although all other procedures would be exactly the same as in the experimental conditions.

Thus we would expect no change in the number of food responses given to

the two sets of slides, for, at the second viewing, *S*'s knowledge that he was part of a control condition would suppress the number of food responses. In this hypothesis we agree with Orne (10), who says that every college sophomore knows that control conditions are not expected to change their responses.

In summary, through the use of a placebo we have arranged two experimental and two control groups. In the control-drug condition we hypothesize that *Ss* will report more food responses to the second set of slides: that is, under the influence of the placebo. In the drug-control condition we hypothesize that *Ss* will report fewer food responses to the second set of slides: that is, without the influence of the placebo. In the two control conditions (control no-pill condition and pure-control condition), we expect that *Ss* will report the same number of food responses to both sets of slides.

C. MATERIALS USED

Each *S* saw a total of 60 slides, 30 under each condition. These slides were bordered with yellow or red tape, so that *S* would be aware (if he was interested) that he was being shown two different sets of slides. Within each set of 30 slides, half were blanks and half were spotted with black smudges. These latter slides looked much as if ink had been dropped randomly over a white sheet of paper, but the blots were not symmetrical. Each *S* was shown two spotted slides and then two blank slides continuing in an AABBA order. Because the spotted slides differed from each other considerably, the order of presentation within a set was kept constant over all *Ss*. Using a slide projector with a Veriac attachment, set at 50 volts and the shutter at .01 second, the slides were projected on a screen about 15 feet away from the *S*. Each slide was shown twice at the .01 speed.

There were two sets of questions: Set A and Set B. Each question was written on an index card for *S*'s convenience in keeping his place, and the *S* wrote his answer on a mimeographed answer sheet. In each set of questions there were 10 multiple-choice questions, 10 loosely structured questions, and 10 projective questions. Each question could be answered with a food response. Examples of the questions are as follows: (a) a multiple-choice question—1) Lemon, sun, Christmas-tree ornament, bracelet. 2) Test tube, mirror, perfume, glass of wine. (b) a loosely structured question—1) A woman. What is she doing? 2) A holiday scene. What is happening? (c) a projective question—1) What is it? (All projective questions were the same.)

Within each set of questions a fixed random order was used. All possible combinations of slides and set of questions were used, each *S* being assigned randomly to any one combination.

D. RESULTS AND DISCUSSION

The placebo effect is one of the most striking results of the experiment and strong evidence for the successful manipulation of the independent variable in at least one of the experimental conditions [see Table 2]. The control-drug condition shows significantly more physical changes as reported on the questionnaire than any of the other conditions. The comparison of this control-

TABLE 2
MEAN NUMBER OF PLACEBO CHANGES IN EACH CONDITION*

Condition	Mean
Control-drug	10.5
Drug-control	7.0
Control no-pill	5.5
Pure control	4.7

* t between control-drug condition and drug-control condition = 2.5, $p = .02$; t between control-drug condition and control no-pill condition = 2.63, $p = .02$; and t between control-drug condition and pure-control condition = 4.22, $p = .01$. No significant differences between the three last conditions.

drug condition with the two control groups that did not receive pills of any kind allow us to allocate the reported physical changes to the demand characteristics of the placebo rather than to the passage of time, to the practice with the questionnaire, or to the 10-minute rest period on the hard table (as many subjects tried to make E believe).

The drug-control condition, however, showed no more placebo changes than did the two control groups and significantly less ($p < .02$) than the control-drug condition; thus it seems that the manipulation was not successful in this order of drug presentation.

A rating made at the conclusion of the experiment offers data for the foregoing hypothesis. Ss were asked to rate their belief in the pill as that belief existed when they were resting immediately after taking the pill (1 = no belief, 7 = belief). No doubt this is a difficult rating to make, for it involves the recollection of a previous state of belief, and Ss may become suspicious about the pill due simply to the fact that the question was asked. Nevertheless, the rapport between Ss and E was unusually good at that time; and we believe that the answers are essentially honest though biased in the upwards direction. Almost one third of the Ss used the rank of seven, the highest rank possible for indicating belief in the pill. Twelve out of 17 Ss ranked their belief as five or higher. Only two Ss used the rank of one, and both of these were in the drug-control condition. When one looks at the ranking made by the control-drug Ss versus those made by the drug-control Ss , he sees that the mean ranking for control-drug Ss is six; while that for drug-control Ss

is only 4.14. The difference is significant at the .01 level. Note should also be taken of the awkwardness of the manipulation of the placebo variable in the drug-control condition: in particular, of the fact that it was necessary to give Ss two pills—one to make them hungry; another to return them to the "full" state. Therefore, any risk involved in insuring the credibility of the placebo is doubled under this condition. On the basis of nonbelief in the pill and awkwardness of the experimental procedure, we conclude that the drug-control condition is an inadequate test of the hypothesis because the situation was not credible to Ss.

In the control-drug condition, Ss showed significantly more food responses under the drug than under the no-drug conditions ($p = .025$, one-tail Walsh test); thus our main hypothesis is confirmed. Demand characteristics can influence S's perception of ambiguous stimuli. No other conditions showed a significant change in the number of food responses over the two viewings of the slides; rather, the number of food responses remained constant within the three other conditions [see Table 3].

TABLE 3
MEAN NUMBER OF FOOD RESPONSES GIVEN BY EACH CONDITION
TO EACH SET OF SLIDES

Condition	Slide Set 1	Slide Set 2	p^*
Control-drug	3.54	5.82	.025
Drug-control	4.3	4.3	n.s.
Control no-pill	3.14	3.85	n.s.
Pure control	4.85	4.21	n.s.

* All tests are one-tailed. The test used is the nonparametric Walsh test (14).

As expected, the drug-control condition appears to be more like a control condition than like an experimental condition. One possible reason for this is the failure of the manipulation of the independent variable. If the pill had worked adequately we would expect that in Ss' first viewing of the slides, which occurred immediately after taking the pill, Ss would respond with a larger number of food items than would those who had not yet taken the pill and, therefore, were still "full." However, there was no difference between conditions on the number of food responses given to the first slide set.

The control no-pill group showed overall the fewest number of food responses and very little change from the first to the second set of slides. On the first set of slides, when they are under the same condition as the control-drug group they show virtually the same number of food responses (3.14 for the control no-pill group *vs.* 3.54 for the control-drug group); however, when they find out that they are in the control condition they do not alter their

level of food responses for the second set of slides. The difference between the control no-pill group and the control-drug group on the second set of slides does not reach the .05 level, but the authors contend that ingestion of the pill and acquiescence to the expectations associated with the pill directly affected perception in the control-drug condition.³

The pure-control group follows the same pattern of no change in the number of food responses given to the two sets of slides. Interestingly, there is a significant difference between the control no-pill group and the pure-control group on the number of food responses given to the first set of slides ($p = .05$). If one assumes that the pure-control group is the baseline group (*Ss* in that group have no knowledge of food, hunger, or pills) while *Ss* in the control no-pill group are food satiated, then the difference in number of food responses present some evidence that the increase in food responses is due not only to the physiological state but also to demand characteristics.

In summary, changes in the perception of blank slides can be influenced by demand characteristics that have been manipulated through the ingestion of a placebo that *Ss* believe will make them hungry. The perceptions do not change if *Ss* receive no pill or if *Ss* do not believe in the pill in spite of the fact that the physiological state of *all Ss* remains unchanged while only the demand characteristics of the experiment vary.

The postexperimental interview had two purposes: to check the credibility of the placebo manipulation and to estimate *Ss*' awareness of the demand characteristics.

Ss were asked to rate their belief in the pill and to describe their own feelings. Many of the physiological changes described by *Ss* would seem serious and disturbing if they were taken out of context, yet none of the *Ss* appeared anxious about the changes. This may be due to their expectations that nothing really serious could happen to them in an experiment. Examples of some of the more serious changes are as follows:

I just feel that I can't see as sharply. The sort of thing as when you're tired, but I don't feel tired. I feel like you feel before you get hungry. Just feel like I'm not focusing. My attention span is lapsing.

As I was lying down I could hear my heart beat and feel my mouth and lips become dry and my head hurt. My right arm got cold and heavy. Can't explain the sensation in my hands, just in my hands, tingling sort of. More aware of noises in the room. Doesn't really seem like a

³ We had no way of measuring the actual physiological state of hunger of each *S*. However since all *Ss* in three of the conditions came to the experiment in a satiated state, we assumed that any change in appetite due to a discussion of food would be distributed randomly over all conditions.

dream but doesn't seem real either. My hands feel like they are shaking though I know they aren't.

As in the second description, many *Ss* complained of a headache because of the 10 minutes spent lying on the hard table; yet not one *S* in the control groups ever attributed a headache to the table.

In general, the placebo reactions were vague and hard to describe, usually centering around feelings of pleasantness, calm, and relaxation. Specific changes, such as moist palms and numb lips, were always similar in topic to the questions that previously had been asked on the first questionnaire. No one ever reported tingling scalp or butterflies in the stomach—topics not touched upon by the questionnaire. Virtually all of the questions in the questionnaire were responded to positively by at least one subject. Only six out of the 46 questions never showed a placebo response. These six questions refer to serious physiological changes, such as feeling ill and abnormal hearing. The most frequent questions to show an effect were those that centered around autonomic-nervous-system reactions pertaining to increased salivation, feeling anxious, and feeling drowsy. Nine or more *Ss* responded positively to these questions.

During the interview each *S* was asked, "Were there more questions pertaining to food on the first set of slides, the second set, or were they about the same?" Regardless of whether the pill was given on the first or second set of slides, 13 out of 17 *Ss* who were asked reported that there were more food questions under the drug condition. There were only three reversals and one report of equal. Certainly, then, *Ss* had food on their mind during the drug period.

Another question in the experiment was, "Now you tell me what do you think is going on in this experiment? What is being tested?" If *S* showed signs of repeating the description previously given by *E*, *S* was urged to report what *he* really thought was being tested. Only one *S* out of 18 was unable to verbalize the purported purpose of the experiment. Some typical answers are as follows:

I knew my mind was supposed to be on food. You told me that the pill would have an effect on my being hungry. If you hadn't said that I would not have seen so much food.

You are trying to find out if by suggesting something it would make me change my perception so that I will see food.

If I had been hungry I would choose appealing food words on the card. If I wasn't I would not.

Thus all *Ss* realized that their physiological state was associated with what

they saw on the screen and that if they were hungry they were supposed to see more food items. Note that this was never mentioned in the instructions or introduction to the experiment.

The demand characteristics of the experiment can never be completely eliminated (10), but investigators must not confuse responses to these demands with responses to the independent variable. In this experiment, the control-drug condition produced results such that if the context of the experiment were unknown they would seem to uphold the Atkinson and McClelland results (3), yet it is clear that the results obtained are artifacts of the experimental design. Even before *Ss* entered the room, they knew that the experiment was connected with food in some way; and they soon discovered that they were to be tested under conditions of satiation and hunger. When *S* first sees the slides it becomes apparent that there are no cues available in the slides themselves; so *S* soon turns to "pure" guessing and the affluent cues available in the experimental procedure. These cues are interpreted by *S* in the following way: Hungry people should think more about food than nonhungry people. When I am in this hungry state I am expected to show more food responses than I did before—when I was not hungry. Although the *Ss* probably do not say this to themselves consciously or clearly, the results show that they react on the basis of this implicit statement.

The two control groups present a different picture. We contend that both of these groups showed no difference in the number of food responses to the two sets of slides because there was no reason to change and no expectation that they should do so. The experimental group (i.e., the control-drug group), on the other hand, had *no more* physiological reason to change than the control groups, but the expectation of the situation was strongly in favor of change, and *Ss* in the group responded to this expectation.

College sophomores or any *Ss* in an experiment are pleased to help *E*, especially if they feel that the experiment is personally important to *E*. At the same time, *Ss* try to assign a meaning to their experience. When both quests can be satisfied by following the cues inherent in the situation, *Ss* will react according to these cues. Such a response is too often quickly and gladly accepted by *E* as confirmation of his hypotheses.

We do not mean to imply that all experimental results are the products of *E*'s implicit wishes, but simply that all experimental designs should be carefully searched for such pitfalls. Perhaps perception is influenced by physiological need and, indeed, it might be; but in the plethora of relevant studies we cannot be sure whether physiological needs influence perception or the need to please *E* influences perception.

REFERENCES

1. ABRAMSON, H. A. Lysergic acid diethylamide (LSD-25): XXX. The questionnaire technique with notes on its use. *J. of Psychol.*, 1960, **49**, 57-63.
2. ABRAMSON, H. A., JARVIK, M. E., LEVINE, A., & KAUFMAN, M. R. Lysergic acid diethylamide (LSD-25): XV. The effects produced by substitution of a tap water placebo. *J. of Psychol.*, 1955, **40**, 367-383.
3. ATKINSON, J. A., & MCCLELLAND, D. C. The projective expression of needs: 1. The effect of different intensities of hunger drive on perception. *J. Exper. Psychol.*, 1948, **25**, 205-223.
4. BARBER, T. X. "Hypnosis," analgesis, and the placebo effect. *J. Amer. Med. Assoc.*, 1960, **172**, 680-683.
5. CONN, J. H. Cultural and clinical aspects of hypnosis, placebos and suggestibility. *Internat. J. Clin. & Exper. Hypn.*, 1959, **7**, 175-185.
6. JACKSON, C. W., & KELLEY, E. L. Influence of suggestion and subject's prior knowledge in research on sensory deprivation. *Science*, 1962, **135**, 211.
7. LAZARUS, R. S., YOUSEM, H., & ARENBERG, D. Hunger and perception. *J. Personal.*, 1953, **21**, 312-328.
8. NEWCOMB, T. M., MURPHY, G., & MURPHY, L. *Experimental Social Psychology*. New York: Harper, 1937.
9. ORNE, M. T. The nature of hypnosis: Artifact or essence. *J. Abn. & Soc. Psychol.*, 1959, **58**, 277-299.
10. ———. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *Amer. Psychologist*, 1962, **17**, 776-784.
11. ORNE, M. T., & SCHEIBE, K. E. The contribution of nondeprivation factors in the production of sensory deprivation effects. *J. Abn. & Soc. Psychol.*, 1964, **68**, 3-12.
12. ROSENBLUM, S., BUONICONTI, P., & GRAHAM, B. D. "Compazine" vs. placebo: A controlled study with educable emotionally disturbed children. *Amer. J. Ment. Defic.*, 1960, **64**, 713-717.
13. SHAPIRO, A. K. A contribution to a history of the placebo effect. *Behav. Sci.*, 1960, **5**, 109-136.
14. SIEGEL, S. *Non-Parametric Statistics*. New York: McGraw-Hill, 1956.
15. WARD, W. D., & SANDVOLD, K. D. Performance expectancy as a determinant of actual performance: A partial replication. *J. Abn. & Soc. Psychol.*, 1963, **67**, 293-295.

Department of Psychology
University of California
Berkeley, California

MEASUREMENT OF RADICALISM-CONSERVATISM*

Department of Psychology, University of California, Los Angeles

ANDREW L. COMREY AND JOHN A. NEWMAYER

A. INTRODUCTION

The isolation of a radicalism-conservatism attitude dimension by factor analytic methods seems to have been reported first by Thurstone (17). Since that time, other investigators (1, 6, 7, 10, 14, 15, 16) have found similar factors. The work of still other investigators leaves some doubt about how general such a factor is or whether or not it is best considered to be a composite of substantially independent variables (8, 9, 13). Most of these studies were carried out at a time when computing facilities placed a severe limitation on the number of variables that could be investigated. This fact, together with the widely varying analytic methods employed and divergent results from previous studies, makes apparent the need for additional research in this area.

The present investigation represents an attempt to apply the factored homogeneous-item-dimension approach (2, 4, 5) to the structuring of political and social attitudes. The essence of this approach is embodied in two principles: (a) knowledge of the factor structure in a domain should be derived from factor analysis of total scores over groups of homogeneous items rather than from factor analysis of single items and (b) groups of items that provide total scores that are alternate forms of each other should be combined into one score rather than included as separate variables in the factor analysis.

Application of the first principle tends to increase the reliability of the variables used; hence favors greater stability of factor results. Application of the second principle avoids the distortion of factor results by overlapping specific variance. Such distortion is common in the factor analysis of single items because it frequently happens that several pairs of similar items appear in the collection to be analyzed. Within such item subgroups of two or even more items there may be substantial item intercorrelations that tend to dominate the factor solution.

B. PROCEDURE

1. *The Sample*

Two hundred and twelve volunteers participated in the study. These individuals were recruited by soliciting university-student volunteers; members of

* Received in the Editorial Office, Provincetown, Massachusetts, on June 17, 1964. Copyright, 1965, by The Journal Press.

Los Angeles social, religious, and fraternal organizations; friends; and volunteers from a form distributed door-to-door in randomly selected blocks in Los Angeles. All participants were promised an analysis of their questionnaire results in return for their cooperation. The subjects may be described in the following terms: Forty-one were under 21, 42 were aged 21 to 24, 25 were 25 to 29, 41 were 30 to 39, 37 were 40 to 54, and 26 were over 54. One hundred and forty of the subjects were male; 72, female. Sixteen subjects had not finished high school, 19 were high-school graduates, 72 had some college work, 57 were college graduates, and 48 had some graduate work. Five subjects described themselves as conservative Republicans, 74 as Republicans, 32 as Independents or as having no political preference, 80 as Democrats, and 12 as liberal Democrats or Socialists.

2. *The Questionnaire*

Thirty political and social-attitude variables were selected for study after a review of previous work and some subjective analysis. An attempt was made to cover major areas without identification with specific issues apt soon to be outdated. An attempt was made to write four homogeneous items to measure each of the 30 attitude variables. Each item was to be answered by a selection from the following scale: 1) Agree Very Strongly, 2) Agree Strongly, 3) Agree, 4) Agree Somewhat, 5) Neither Agree nor Disagree, 6) Disagree Somewhat, 7) Disagree, 8) Disagree Strongly, 9) Disagree Very Strongly. The 120 items were arranged in a booklet—one item for each dimension in the first 30 items, a second item for each dimension in the second 30 items, and so on. Subjects recorded their responses on a separate answer sheet, writing merely the number of the response selected. These numbers were treated as scores.

Scores for the 120 items were intercorrelated and factor analyzed for item-analysis purposes. Factors were extracted by the minimum-residual method (3), and rotated orthogonally by the normal-varimax criterion (12). Inspection of the factor results and the intercorrelations among items reveals the following facts:

1. All the items conceived to measure Religionism and Fundamentalism showed up on the same item factor, with loadings in excess of .60; and intercorrelations across item groups were as high as those within groups. Therefore, the items for these two groups were combined to form one dimension: Religiosity.

2. Items for two variables originally called Pacifism and Militarism also combined to form the same item factor, albeit at opposite ends of the continuum. Seven items of the original eight items had loadings of .68 or more on this item factor; hence were combined to form one dimension: Pacifism. The militarism items were reversed in the scoring, of course.

3. Items for the variables Weak Federal Government and Pro-Free-Enterprise defined a single item factor, all eight items having loadings of .46 or more. These items were combined to measure one dimension: Weak Federal Government.

4. Items for the variables Censorship and Restrictions on Sexual Freedom were combined for the dimension Moral Censorship, although the variance (for all items) for these two variables divided between two item factors.

5. The variable Birth Control split into two distinct item factors (two items each): Population Control and Contraception.

6. As expected, several other variables showed up as item factors and were retained as dimensions; although in many cases less than four items proved to have factor loadings high enough for the items to be retained. These dimensions were Antiunionism, Pay Based on Ability, Partisanship, World Government, Fatalism, Belief in People, and Education for Adjustment.

7. Several other variables failed to define separate item factors but appeared on item factors with other groups of items. Inspection of item factor loadings, patterns of interitem correlations, and item content, however, failed to justify pooling these items with others but suggested their retention as separate variables. These dimensions were Welfarism, Racial Tolerance, Severe Treatment of Criminals, Service to Country, Rapid Social Change, Traditional Moral Values, Fascism, Respect for Age, Service to the Individual, Indoctrination, and Cultural Ethnocentrism.

Twenty-nine variables were selected for the main factor analysis. Twenty-five of these were total scores over groups of two or more homogeneous items. The last four variables consisted of single biographical items. The names of the variables and the items used to measure them are presented later. After each variable name, the mean, standard deviation, and a reliability estimate for that variable is given. The reliability estimates were obtained by computing the average interitem correlation to obtain the reliability of a test of the appropriate number of items. The A and B and the number following certain items represent the form to which the item has been assigned and the factor loading for the item. This will be explained further later on.

C. ITEMS

1. *Religiosity* (47.8, 15.3, .94)

Every child should have religious instruction (A, .70).

God exists, in the form in which the Bible describes him (A, .67).

This country would be better off if religion had a greater influence in daily life (A, .62).

School teachers should believe in God (B, .73).

It should be against the law to do anything which the Bible says is wrong (B, .59).

Moses got the ten commandments directly from God (B, .66).

All people alive today are the descendants of Adam and Eve (A, .53).

All the miracles described in the Bible really happened (B, .59).

2. *Pacifism* (45.2, 13.5, .93)

Under no circumstances should our country use nuclear bombs against anybody (B, —.56).

I would rather have a foreign power take over our country than start another world war to stop it (B, —.70).

This country should disarm regardless of whether or not other countries do (A, —.66).

If my country had been destroyed, I still would not push the button to wipe out the attacking enemy nation (A, —.62).

Our military forces should be kept strong enough to defeat any nation that might attack us or our allies.

Our country should be engaged constantly in research to develop superior weapons for our national defense (A, .66).

Our country should prepare to employ every available weapon to destroy any major power that seriously attacks us (B, .73).

3. *Welfarism* (20.6, 7.4, .82)

The average man today is getting less than his rightful share of our national wealth (A, —.52).

The government should guarantee every citizen enough to eat (A, —.47).

It is the responsibility of the government to take care of people who can't take care of themselves (B, —.43).

If the government must go deeper in debt to help people, it should do so (B, —.66).

4. *Anti-Unionism* (20.6, 5.8, .77)

Many large unions have officers with criminal records (A, .38).

Most unions try to prevent the efficient use of labor (B, .40).

Most unions do not elect officers by honest, secret-ballot elections (A, .46).

Many union leaders use threats and violence to keep themselves in power (B, .46).

5. *Weak Federal Government* (36.4, 13.8, .92)

Central government should run only those things which cannot be run effectively at the local level (A, .52).

The federal government has too much power over citizens and local government (A, .62).

Greater decentralization of power would be better for this country (A, .64).

The federal government should not interfere in the affairs of individual states unless absolutely necessary (B, .66).

The strength of this country today is largely a product of the free enterprise system (B, .68).

A greater degree of government control over business would result in a weakening of this country's economy (A, .73).

Regulation of business by government usually does more harm than good (B, .65).

When something is run by the government, it is apt to be inefficient and wasteful (B, .47).

6. *Moral Censorship* (55.3, 13.1, .90)

Motion pictures which offend any sizeable religious group should be banned (B, .49).

If a man is showing a sex movie to friends in his own home, the police should stop it (A, .43).

Every city should prevent the sale of objectionable books (A, .68).

Public libraries should contain only books which are morally sound (B, .62).

Sexual relations between unmarried people should be illegal (A, .60).

The police should hunt down homosexuals and put them in jail (A, .61).

A woman who has sexual relations with a man for money should go to jail (B, .58).

More restrictions should be imposed to prevent young people from having sexual relations before marriage (B, .68).

7. *Population Control* (5.3, 3.0, .63)

In overpopulated countries the people should be taught to practice birth control.

Uncontrolled population growth is a dangerous menace.

8. *Contraception* (5.8, 3.6, .62)

Abortion should be legalized (A, —.47).

Birth control devices should be made readily available to anyone who wants to use them (B, —.46).

9. *Racial Tolerance* (13.4, 5.6, .68)

Marriages between persons of different races should be socially acceptable (B, —.67).

Employers should be prevented by law from hiring only people of their own race (A, —.39).

Social organizations limiting membership to a particular race should be abolished.

10. *Severe Treatment of Criminals* (21.7, 6.4, .82)

Teenage hoodlums should be punished severely (B, .59).

Our laws give too much protection to criminals (B, .60).

Criminals convicted of three separate felonies should never be released (A, .55).

In our country, the sentences handed out to criminals are usually too light (A, .66).

11. *Pay Based on Ability* (11.1, 3.8, .62)

A well educated person should have a higher income than one who is less well educated.

Jobs requiring a lot of brains should pay much more than those which do not.

A person who works hard should earn more than a coworker who doesn't work hard.

12. *Capital Punishment* (20.3, 8.1, .86)

A mentally ill man who attacks and kills a little girl should be executed (A, .57).

A dictator who orders the extermination of thousands of innocent people should be executed for his crimes (B, .57).

A gunman who kills someone in an armed robbery should receive the death sentence (A, .69).

Someone who plans and carries out the murder of his or her spouse should be executed (B, .71).

13. *Partisanship* (26.1, 4.6, .68)

The political party which I favor is unquestionably better for the country than any other present party.

I would not support a political candidate of the major opposition party to my own.

Leaders of the major political party opposite to my own party are generally cynical, selfish men who disregard the national welfare.

I am inclined to dislike anyone who can accept the view of the major political party opposite to my own.

14. *Service to Country* (18.0, 7.6, .85)

Every able bodied male should willingly serve for a period of time in his country's military service (A, .70).

If called upon to do so, a citizen should be willing to sacrifice his life for his country (B, .59).

Patriotism is one of the great virtues (B, .74).

A man who is ready to die for his country deserves the highest honor (A, .62).

15. *Rapid Social Change* (15.7, 4.7, .65)

Where social injustice exists, it must be changed immediately regardless of the consequences.

People who are not willing to give others equal rights should be required by law to do so.

People who discriminate against others because of race or religion should be treated like criminals.

16. *World Government* (19.8, 9.0, .90)

The United States should work peacefully for a strong world government (A, —.55).

The United States should be willing to surrender some of its rights to strengthen the United Nations (A, —.80).

The United States eventually should give up its military power to a strong world government (B, —.74).

Present nations should become states within an all powerful world government (B, —.61).

17. *Fatalism* (15.2, 2.8, .67)

Man cannot alter his destiny.

Each person's life is laid out for him in advance.

18. *Traditional Moral Values* (6.0, 2.8, .52)

People should work hard and pay their debts on time.

A man should always be faithful to his wife.

19. *Fascism* (11.8, 3.3, .52)

A nation will generally progress more rapidly under the guidance of a single strong, benevolent leader.

Representative bodies like Congress have too many arguments and divisions to function effectively.

20. *Respect for Age* (10.2, 3.1, .63)

As people get older, they usually become wiser.

Older people usually have a more intelligent view of life than younger people.

21. *Service to the Individual* (10.3, 3.6, .73)

Laws which benefit the people are more important than laws which strengthen the nation (A, —.63).

The welfare of the individual is more important than the welfare of the country (B, —.48).

22. *Indoctrination* (18.3, 6.6, .77)

Children should be taught early that our country is the greatest country in the world.

There are things a child must be taught to believe before he is ready to understand them.

Moral values should be instilled in children while they are very young.

Children should be taught to obey their parents without question or argument.

23. *Belief in People* (13.0, 3.9, .68)

The average person, if suddenly handed a government position, would try to do an honest and hardworking job.

Only a few people would deliberately lie to get welfare support money from the government.

Most people are basically honest.

24. *Cultural Ethnocentrism* (29.3, 4.4, .73)

Cultural exchanges with hostile nations are of little benefit to us.

The people of this nation have little to learn from peoples of other nations.

Americans should patronize American art and music rather than foreign imports.

American students should be shielded from foreign ideologies and philosophies until they are adults.

25. *Education for Adjustment* (11.0, 3.8, .73)

Schools should spend more time teaching children how to get along with each other.

High school education should be more concerned with helping the student make a good social adjustment.

26. *Age* (2.3, 1.7, —)

Below 21 was scored as 0; 21-24, 1; 25-29, 2; 30-39, 3; 40-54, 4; and over 54, 5.

27. *Sex* (0.4, 0.5, —)

Males were scored 0; females were scored 1.

28. *Educational Level* (2.5, 1.2, —)

Less than high school graduation was scored 0; high school graduate, 1; some college, 2; college graduate, 3; some graduate work, 4.

29. *Political Preference* (2.1, 1.0, —)

Conservative Republicans were scored 0; Republican, 1; Independent or no preference, 2; Democrat, 3; liberal Democrat or Socialist, 4.

D. THE ANALYSIS

A total score was computed for each of the 212 subjects over each of the 25 item groups by adding the numbers of the responses selected by the subject for those items. In the cases of Pacifism, the scores for three of the items had to be subtracted from 10 before adding to maintain a uniform direction in the items scored because these three items measured the opposite end of the continuum. These 25 scores, plus scores for age, sex, educational level, and political preference were intercorrelated (Pearson r). This 29-by-29 correlation matrix was factor analyzed by the minimum-residual method (3), and nine factors were extracted. This method is similar to the principal-components method of factor analysis except that no communality estimates are used or needed. Convergence could not be obtained for a tenth factor.

The nine factors were rotated orthogonally by the normal-varimax criterion (12). These orthogonal factors were further rotated obliquely using simple structure and psychological meaningfulness as criteria to guide the rotations. The reference-vector projections and the correlations among reference vectors were utilized to obtain the table of correlations among the factors.¹ The table of correlations among the factors was analyzed by the principal-factor method, using unities in the diagonals, to provide knowledge of the second-order factor domain.

E. RESULTS

Five of the nine factors had at least two factor loadings over .60. These factors will be presented along with the variables that had factor loadings of .30 or more.

1. *Factor II. Welfare-State Attitudes*

3. Welfarism, —.82; 29. Political Preference, .78; 5. Weak Federal Government, .68; 4. AntiUnionism, .50; 23. Worthiness of People, —.42; 15. Rapid Social Change, —.39; and 21. Service to the Individual, —.32.

2. *Factor III. Punitive Attitudes*

10. Severe Treatment of Criminals, —.70; 12. Capital Punishment, —.66; 6. Moral Censorship, —.47; 7. Population Control, —.46; 19. Fasc-

¹ The computations were carried out on the IBM 7090, operated by the UCLA computing facility. The following items have been deposited with the ADI: the matrix of correlations among dimensions, the minimum residual factor matrix, the varimax factor matrix, the oblique reference vector projection matrix and the transformation matrix to get it from the varimax matrix, the oblique factor pattern, and the oblique factor structure. Order Document No. 8507 from the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., remitting in advance \$1.25 for 35 mm microfilm or \$1.25 for photocopies. Make check payable to Chief, Photoduplication Service, Library of Congress.

ism, —.37; 20. Respect for Age, .37; and 24. Cultural Ethnocentrism, —.33.

3. Factor IV. Nationalism

14. Service to Country, —.83; 2. Pacifism, .82; 16. World Government, .66; 7. Population Control, .49; 21. Service to the Individual, .48; 22. Indoctrination, —.43; 19. Fascism, .31; and Age, .30.

4. Factor VII. Religious Attitudes

1. General Religiosity, —.69; 8. Contraception, .76; 6. Moral Censorship, —.65; 7. Population Control, .44; 24. Cultural Ethnocentrism, —.34; and 17. Fatalism, —.32.

5. Factor IX. Racial Tolerance Attitudes

15. Rapid Social Change, —.70; Racial Tolerance, —.65; and 25. Education for Adjustment, —.31.

The negative loadings for the attitude variables are negative when consistent with the factor-title direction due to the scoring of "1" for Agree Very Strongly and "9" for Disagree Very Strongly.

The remaining four factors are of lesser importance. Factor I had loadings of .46, .38, and .33 for variables 20, 13, and 27, respectively. Factor V had loadings of .60 and .34 for variables 26 and 27. Factor VI had loadings of .56, .47, and .38 for variables 18, 11, and 21 respectively. Factor VIII had loadings of .49 and .40 for variables 25 and 28. The intercorrelations among factors are given in Table 1.

TABLE 1
CORRELATIONS AMONG FACTORS*

Factor	1	2	3	4	Factor	5	6	7	8	9
1										
2	—10									
3	32	—56								
4	00	69	—71							
5	—05	36	—29	35						
6	16	—35	41	—45	—16					
7	04	—48	46	—60	—23	26				
8	02	—32	42	—48	—28	12	29			
9	16	55	—43	52	22	—10	—31	—23		

* See p. 358 for explanation of factors.

The matrix of correlations among the factors, shown in Table 1, was factor analyzed using unities in the diagonals. Only one major second-order

factor resulted, accounting for 42 per cent of the total variance. Loadings for the five most-important primary factors on this second order factor were as follows: II. Welfare State Attitudes, .81; III. Punitive Attitudes, —.82; IV. Nationalism, —.90; VII. Religious Attitudes, —.68; and IX. Racial Tolerance Attitudes, .62. Loadings for factors I, V, VI, and VIII were —.14, .49, —.51, and .56 respectively. This second-order factor is considered to represent the radicalism-conservatism attitude continuum. At the radical end of the continuum we have belief in the welfare state, powerful federal government, pacifism, world government, racial tolerance, and rapid social change. On the conservative end of the continuum, we have belief in severe treatment of criminals, capital punishment, religion, moral censorship, and service to country.

Dimensions 1 through 10, 12, 14, 15, 16, and 21 had substantial loadings on the five primary factors that best defined the second-order factor of radicalism-conservatism. These 15 dimensions were measured by a total of 67 items. To select items for measuring radicalism-conservatism, these 67 items were intercorrelated (by Pearson r), and a single minimum-residual factor extracted. Thirty pairs of matched items were selected from the total of 67 items to provide two forms, A and B, of a radicalism-conservatism test. The number of items taken from any given dimension was the same in both forms. Items in the two forms were also matched for size of loadings on the single general factor in the analysis of 67 items. Those items selected for Form A are those which have an "A" in parentheses following the item as shown above. The factor loading on the general factor from the 67-item analysis follows the A. Items for Form B and their factor loadings are similarly indicated. If the same scale of response is used as that employed here, scores on the items stated in the radicalism direction must be reversed by subtracting each such item score from 10. This is required for each item that has a negative factor loading after the letter A or B above. Adding up item scores, including the reversed item scores, for the 30 items in Form A and the 30 items in Form B will give two measures of radicalism. The low scores indicate conservatism.

In the present sample, for Form A, the mean and standard deviation, respectively, were 159.0 and 41.3. For Form B, they were 158.4 and 41.8 respectively. For the total score, Form A plus Form B, the corresponding statistics were 317.4 and 82.3. The alternate-forms reliability estimate is .96, correlating Form A and Form B. This corrects to .98 for the total score, although these figures can be expected to drop with a new sample. The distributions are approximately normal in this sample.

F. DISCUSSION

This investigation helps to resolve some of the differences that have appeared between findings in Great Britain and this country. Ferguson (8, 9), for example, structured the domain in terms of three primary social attitudes, Religionism, Humanitarianism, and Nationalism, with no Radicalism-Conservatism factor. His three primary social attitudes correspond closely to three of the primary factors obtained here, Religious Attitudes, Punitive Attitudes, and Nationalism respectively. The Radicalism-Conservatism factor is to be found in the substantial correlations between his three primary social attitudes. In fact, because of this correlation, Ferguson found it necessary to analyze two attitudes at a time rather than all three together.

British authors, on the other hand, have typically extracted few factors and have more often than not left them unrotated. If there is a general factor, as in the present case, it will appear at this first unrotated factor. Thus, the British authors have identified Radicalism-Conservatism as a first unrotated factor.

Thus, rotation of first-order factors to either orthogonal or oblique positions approximating simple structure tends to organize the attitude domain around several distinct variables (such as religionism, nationalism, humanitarianism, and so on) with no Radicalism-Conservatism factor. These same variables factor analyzed without rotation will show Radicalism-Conservatism as the first general factor, but will not reveal appropriately the correlated primary factor-level components of Radicalism-Conservatism. The present analysis, which uses oblique rotation to the correlated primary factors, and then second-order analysis is the appropriate means for demonstrating the existence of both the general factor and its correlated primary components.

REFERENCES

1. CARLSON, L. A factorial study of attitudes. *J. Soc. Psychol.*, 1934, 5, 202-212.
2. COMREY, A. L. Factored homogeneous item dimensions in personality research. *Educ. & Psychol. Meas.*, 1961, 21, 417-431.
3. ———. The minimum residual method of factor analysis. *Psychol. Reports*, 1962, 11, 15-18.
4. ———. A study of thirty-five personality dimensions. *Educ. & Psychol. Meas.*, 1962, 22, 543-552.
5. COMREY, A. L., & SCHLESINGER, B. Verification and extension of a system of personality dimensions. *J. Appl. Psychol.*, 1962, 46, 257-262.
6. EYSENCK, H. J. General social attitudes. *J. Soc. Psychol.*, 1944, 19, 202-227.
7. ———. *The Structure of Human Personality*. London: Methuen, 1960.
8. FERGUSON, L. W. The measurement of primary social attitudes. *J. of Psychol.*, 1940, 10, 199-205.

9. ———. The isolation and measurement of nationalism. *J. Soc. Psychol.*, 1942, **16**, 215-228.
10. FRUCHTER, B., ROKEACH, M., & NOVAK, E. G. A factorial study of dogmatism, opinionation, and related scales. *Psychol. Reports*, 1958, **4**, 19-22.
11. HARMAN, H. H. *Modern Factor Analysis*. Chicago: Univ. Chicago Press, 1960.
12. KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, **23**, 187-200.
13. KERR, W. A. Untangling the liberalism-conservatism continuum. *J. Soc. Psychol.*, 1952, **35**, 111-125.
14. LORR, M. A factorial isolation of two social attitudes. *J. Soc. Psychol.*, 1951, **34**, 139-142.
15. SANAI, M. A factorial study of social attitudes. *J. Soc. Psychol.*, 1950, **31**, 167-182.
16. ———. An experimental study of social attitudes. *J. Soc. Psychol.*, 1951, **34**, 235-264.
17. THURSTONE, L. L. The vectors of mind. *Psychol. Rev.*, 1934, **41**, 1-32.

Department of Psychology
University of California
Los Angeles, California 90024

THE LANGUAGE OF POSTURAL INTERPRETATION*¹

Department of Psychology, Wesleyan University

ROBERT H. KNAPP

A. INTRODUCTION

In earlier papers (1, 2, 4), the author has discussed the "diadic silhouette." One of those papers reports the distinctive manner in which male schizophrenics are prone to interpret such stimuli. Another deals with how the judged acceptance of one figure by the other is influenced by alterations in posture. A third describes the manner in which different types of interpersonal transactions are ascribed to different versions of the diadic silhouette. In the present article, the author reports data roughly parallel to those in the third of the previous articles. This article differs from the earlier one in that it presents a level of generality not reported heretofore, it deals with a new and improved series of stimuli, and it calls for a pattern of responses altered from that presented in the earlier studies.

In the Diadic Silhouette Test, two profile human figures (male and female) are presented in 27 postural variations. Nine of the postural variations involve a condition in which the female is at a higher elevation than the male, nine in which the two figures are at equal levels, and nine in which the male figure is higher. Further variation is introduced by rotating the male and female figures from backward to upright to forward positions. The combinations of three elevation patterns with the three rotational positions for each figure yield 27 silhouette variations. A departure from the stimuli used in earlier studies (1, 2) lies in the fact that the figures in the present instance were mature and not juvenile, and three rotations of the figure rather than four were employed.

The subjects were required to view each diadic silhouette and to describe the nature of the transaction or exchange occurring between the two figures. For that purpose they were provided with a "menu" consisting of 16 items. They were asked to specify which of the two figures was initiating the exchange, whether the "act" involved was "telling," "asking," "demanding," or "offering," and whether the other party was responding affirmatively or nega-

* Received in the Editorial Office, Provincetown, Massachusetts, on June 26, 1964. Copyright, 1965, by The Journal Press.

¹ Support for this research was obtained in part from the National Institutes of Mental Health under Contract M-2178 to Wesleyan University.

tively to the initiated action. In this study the "menu" was identical with that in the previous studies except that the act "demanding" replaced "commanding." It should be observed in passing that each subject was free to employ each of the 16 menu items as often as he felt it appropriate.

B. METHOD AND PROCEDURE

The subjects consisted of 75 adult teachers (38 women and 37 men) engaged in summer-school training. They were presented with the 27 diadic silhouettes in two forms: first, with a bluish-green background; second, with a reddish background. The differences in interpretation associated with the color of the background proved to be comparatively small and will not be discussed. One should note, however, that each individual rated the 27 silhouettes twice, once in its blue and again in its red form, thus yielding two responses to each silhouette.

Just as the color of the background proved to be a relatively unimportant determinant of the interpretation placed on the silhouette; so sex differences (though persistent and significant) were not overwhelming. These differences, which have been reported elsewhere (3) are more associated with the propensity to employ different items of the menu than with differences in the interpretation of particular silhouettes.

C. RESULTS

The manner in which each of the 16 items of the menu can be associated with the patterns of elevation and rotation have been given earlier (2). Here the author presents more generalized results, based upon the revised and improved form of the Diadic Silhouette Test. Tables 1 to 6 show how each of the component elements of the menu is ascribed to the 27 diad variants. Table 1 shows the incidence at which, for each silhouette, the male figure is seen as the initiator. For the tables we have calculated chi square with 26 degrees of freedom and based upon the null hypothesis that a response is equally likely to occur in any of the 27 cells. Accordingly we are establishing the fact that the incidence of responses in the different cells departs significantly from randomness and that these departures possess an internal coherence. On the null hypothesis we should expect 79 responses in each cell of Table 1. The χ^2 of 243 exceeds the .01 level of confidence.

There is a fairly incisive pattern for all three elevations. In Elevation 1 it appears that the male is seen as the initiator most frequently (when both parties have their heads thrown backward), and very rarely are both heads rotated forward. For this elevation pattern, the male is described as the

TABLE 1
THE INCIDENCES OF RESPONSES INVOLVING THE MALE INITIATOR
FOR EACH OF THE 27 DIADIC VARIANTS*

Female	Elevation 1 (Female higher)			Male Elevation 2 (Equal)			Elevation 3 (Male higher)		
	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward
Backward	103	80	70	78	99	84	37	71	107
Upright	65	56	44	70	103	68	71	86	129
Forward	56	34	39	64	96	99	83	100	140

* $\chi^2 = 243.3$. Each subject contributed two responses to each silhouette: first, when the background was red; second, when it was blue. In computing chi square we have combined all responses of all subjects. This constitutes a violation of the requirements of independence for chi square; however, the chi-square values as computed are except one instance far in excess of the .01 level of confidence. Moreover, the case for the significance of our findings lies not upon the value of chi square itself but upon the coherence of the emerging pattern for each of the variables examined.

initiator decreasingly as either figure moves his head in a forward direction; though it is suggested that the rotation of the female figure exercises a greater determinant than that of the male.

In the second elevation position (both figures equal) the pattern is less clear, and it probably must suffice for us to observe that the male is seen as the initiator when his head is upright. The pattern in Elevation 3 (male higher) is virtually the opposite of that obtaining in Elevation 1. Again, it appears that the rotation of both figures is of significant importance and that, as the rotation is forward, the male figure is less often seen as the initiator. But the position of the male figure now appears to be a more important determinant than that of the female.

Let us now consider which silhouettes tend to evoke interpretations involving the act of "telling." The incidence with which such ascriptions are made is found in Table 2. It is evident that the difference between the three

TABLE 2
THE INCIDENCES OF RESPONSES INVOLVING "TELLING" FOR
EACH OF THE 27 DIADIC VARIANTS*

Female	Elevation 1 (Female higher)			Male Elevation 2 (Equal)			Elevation 3 (Male higher)		
	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward
Backward	34	43	49	37	59	50	47	59	42
Upright	51	90	67	55	95	59	44	72	44
Forward	40	68	45	36	57	66	39	51	52

* $\chi^2 = 110.87$. See footnote to Table 1.

elevation positions is relatively slight, but that in all three a conspicuously heightened incidence is observed when the two figures are upright, looking directly forward. In this instance the null hypothesis leads one to expect 54 in each cell. The obtained chi square of 110 exceeds the .01 level of confidence.

TABLE 3
THE INCIDENCES OF RESPONSES INVOLVING "ASKING" FOR
EACH OF THE 27 DIADIC VARIANTS*

Female	Elevation 1 (Female higher)			Male Elevation 2 (Equal)			Elevation 3 (Male higher)		
	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward
Backward	42	42	38	19	41	36	66	47	38
Upright	36	28	38	51	25	39	46	44	42
Forward	49	41	37	46	39	35	34	34	28

* $\chi^2 = 54.59$. See footnote to Table 1.

Table 3 presents the incidence at which transactions involving "asking" are ascribed to each of the silhouettes. Here the null expectation is 39 per cell. The obtained chi square of 59 greatly exceeds the .01 level of confidence. The pattern is less incisive, but still is clear for the first two elevation patterns at least. There appears to be a hollow-center effect such that "asking is significantly more rarely seen where both figures are in upright rotation. Backward or forward rotation of either figure is more likely to elicit this type of interpretation. Elevation 3 yields a less-coherent pattern, though there appears to be a kind of diagonal gradient. When both heads are backward, a high incidence of "asking" is elicited; but when both heads are forward this interpretation is infrequent. On balance, the hollow-center tendency is probably the most important to be noted.

Table 4 shows an incisive pattern for "demanding," which stands in direct

TABLE 4
THE INCIDENCES OF RESPONSES INVOLVING "DEMANDING"
FOR EACH OF THE 27 DIADIC VARIANTS*

Female	Elevation 1 (Female higher)			Male Elevation 2 (Equal)			Elevation 3 (Male higher)		
	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward
Backward	47	43	42	74	27	42	21	19	37
Upright	37	9	22	20	12	31	35	12	42
Forward	39	20	46	42	33	22	47	34	39

* $\chi^2 = 157.36$. See footnote to Table 1.

opposition to that for the act of "telling." We might expect this because "telling" is the most impersonal of exchanges, whereas "demanding" is the most personal and imperative. Here we find the hollow-center effect startlingly visible. Our null expectation for each cell is 33. The chi square of 157 attains a highly significant value. It will be observed that for all three elevation patterns the center cell shows a very low incidence (nine, 12, and 12, respectively). It is clear that this highly charged type of interpersonal transaction is rarely seen when the figures are both in upright rotation. On the contrary, this interpretation is most notably applied when the figures are on the same level (each with his head thrown back), a position that is commonly interpreted as mutual defiance.

In Table 5 we consider the fourth act, "offering." It is clear that the

TABLE 5
THE INCIDENCES OF RESPONSES INVOLVING "OFFERING" FOR
EACH OF THE 27 DIADIC VARIANTS*

Female	Elevation 1 (Female higher)			Male Elevation 2 (Equal)			Elevation 3 (Male higher)		
	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward
Backward	27	21	21	16	22	22	14	24	33
Upright	26	23	21	24	16	20	27	22	22
Forward	22	20	22	25	21	27	30	32	31

* $\chi^2 = 21.33$. See footnote to Table 1.

pattern here is far less incisive than for the components reviewed earlier. We would expect on the null hypothesis 23 responses per cell. The chi square for this table is but 21 and without statistical reliability. There is lacking here any manifest or significant pattern, such as observed for other components of the "menu."

Finally, as may be seen in Table 6, we have calculated the frequency of

TABLE 6
THE INCIDENCES OF RESPONSES INVOLVING POSITIVE RESPONSES
FOR EACH OF THE 27 DIADIC VARIANTS*

Female	Elevation 1 (Female higher)			Male Elevation 2 (Equal)			Elevation 3 (Male higher)		
	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward	Back- ward	Up- right	For- ward
Backward	59	63	76	53	77	68	76	110	114
Upright	96	129	104	83	122	116	75	108	113
Forward	104	126	99	67	113	121	71	99	103

* $\chi^2 = 142.29$. See footnote to Table 1.

positive responses for each of the 27 silhouettes independently of the initiator or the act involved. On the null hypothesis we would expect 94 responses per cell. The obtained chi square of 142 greatly exceeds the .01 level of confidence.

The pattern here turns out to be entirely lucid. In Elevation 1 in which the female figure is higher, the incidence of positive responses is reduced when the head of the female figure is rotated backwards. In Elevation 2 it is apparent that if either figure is rotated backwards, the incidence is conspicuously low. In Elevation 3, it is clearly manifest that the backward rotation of the male figure (now at a higher elevation) is associated with a low incidence of positive responses. Thus, it is clear that the most-important determinant in Elevations 1 and 3 is the rotation of the most elevated figure. In Elevation 2 the rotation of both heads appears to exert an approximately equal influence on the interpretation. The relationship revealed strikes the author as a particularly convincing demonstration of the importance of the rotation of the figures in interaction, with elevation pattern as a determinant of the subject's interpretations.

D. DISCUSSION

The present article does not seek to illuminate individual differences in the perception of diadic silhouettes. It should be thought considered as a contribution to that unexplored language of postural relations so central to the art of the theater and other visual arts involving the depiction of human figure, such as portraiture, iconography, and sculpture. The author has presented, he believes, convincing evidence that certain dynamic human relationships are best conveyed by certain postural attitudes of the figures involved. Moreover, the relation of these two can be reduced to some rough and approximate lawfulness. In fact, in all cases except the perception of relations involving "offering" and possibly "asking" highly coherent results have been obtained. These results can be defended not merely through an appeal to statistical significance but more convincingly by the logical coherence that they incorporate.

It is interesting that in the judgments reported (which are essentially psychophysical), there repeatedly recur interactions between all three of the variants of the stimulus and the imposed interpretations; for example, the incidence of male initiation is clearly related to all three in a complex interactive sense. The same may be said for the incidence of positive responses. On the other hand, responses involving "telling" and "commanding" appear to be relatively independent of elevation but strongly determined by the interaction of the two head rotations.

REFERENCES

1. KNAPP, R. H. Perceptual interpretation of the diad: I. Judgments of acceptance. *J. Soc. Psychol.*, 1964, **63**, 169-177.
2. ———. Perceptual interpretation of the diad: II. Character of relationship. *J. Soc. Psychol.*, 1964, **64**, 89-100.
3. KNAPP, R. H., & EHLINGER, L. A. Sex differences among schizophrenics in the interpretation of the human diad. *J. of Psychol.*, 1963, **56**, 473-478.
4. ———. Sex differences in the incidences of responses to the Diadic Silhouette. Test. *J. Soc. Psychol.*, 1966, **68**, in press.

Department of Psychology
Wesleyan University
Middletown, Connecticut

MODES OF CONFORMITY AND CAREER SELECTION OF RURAL AND URBAN HIGH-SCHOOL SENIORS*

State University of New York at Buffalo

EUGENE L. GAIER AND WILLIAM F. WHITE¹

A. INTRODUCTION

The senior year of high school represents for many the transition from passive dependency to a state of conflict where the student is forced to consider decisions concerning career choices as well as other patterns of future activity. This period may merely represent the formalization of four years of preparation or it may become the precipitating moment for choice as the secondary school years close their doors. But career decisions are usually not made randomly. The socialization process to which the individual is exposed early starts to exert pressures concerning career and occupational choices he will eventually make and:

. . . The occupations among which an individual makes his final choices do not usually vary very much. It is not certain that they always vary sufficiently to form the hierarchy presupposed by the level of aspiration theory. For example, the hierarchy used by Svalastoga in his empirical study consists of occupations ranging from shoe-shine boy to ambassador. It is most unlikely that anybody ever has to make his decision subject to such an extensive range of variation. (2, pp. 240.)

When Reisman (6, 7) hypothesized his trichotomy of character orientation—tradition directed, inner-directed and other-directed—he also sought its application in the economy wherein the individual was socialized. As pointed out by Littunen and Gaier (5):

Reisman's tradition directed personality belongs to the agrarian, pre-industrial economy, the inner-directed personality appears to emerge in the context of an expanding, industrializing economy, while the other-directed personality type represents the highly industrialized economy of abundance. . . . whereas, the inner-directed personality conforms to the personal ideal laid down in childhood by his parents and the older generation in general, the other-directed personality usually conforms to his peer group. By extension, an economy of abundance and his industrialization fosters larger numbers of other-directed personalities.

* Received in the Editorial Office, Provincetown, Massachusetts, on July 6, 1964. Copyright, 1965, by The Journal Press.

¹ Now at the University of Georgia.

B. THE PROBLEM

The purpose of the present study was to examine the cultural determinants toward educational and occupational goals of two senior classes: one in an urban setting; the other, in a rural context. In this study, we were not as interested in the content of the conforming influence as we were in the mechanism of conformity and its influence on decisions toward careers and occupations. Specifically, we were interested in assessing the direction of the anomie-automatic axis that the senior high-school student may move. For example, does the person who has been strongly conformed by inner-directed culture show a significant direction toward education? Has the student's character, substantially weighted with an other-directed environment, been socially set toward anomie or autonomy? Does other-directed culture conformity bear any relationship to the desire for higher education?

C. SUBJECTS AND PROCEDURE

Eighty-six graduating seniors in a high school situated in a large urban community and 84 graduating seniors in a small rural community were randomly selected from their respective classes and asked to list up to three of each of what they considered their greatest personality assets and personality liabilities about themselves, their mothers, and their fathers. In addition, each subject was asked to complete a number of questions concerning his future plans. The results reported in the present paper are based on a sample of 170 seniors (97 males and 73 females) ranging in age from 16 to 20.

The questionnaire, adapted from studies reported by Reisman (5, 6) and revised in earlier investigations reported by Gaier and Wambach (4) and Littunen and Gaier (5), was administered in the respective schools during class hours.

D. RESULTS AND DISCUSSION

1. Assets

The 1,386 personality assets obtained were grouped into three major areas: Group behavior, Character traits, and Achievement, each including a number of subcategories.¹ Typical asset responses classified as group behavior are presented below, together with the subcategories within each breakdown:

- a. *Getting along with others.* "The ability to get along O.K. with other people."
- b. *Adjusting to the environment.* "She can face any crisis."
- c. *Concern for people.* "She is always willing to help others."
- d. *Friendships.* "My ability to make friends."

¹ See Gaier and Wambach (4) for a full treatment of the rationale of the classification scheme developed.

Positive character traits emphasized traditional virtues and typically included:

- a. *Striving*. "His strong desire to be somebody."
- b. *Character Strength*. "His general good moral standards."

Typical responses expressed in the four subcategories of the area of achievement include the following:

- a. *Intellectual*. "My ability in school subjects."
- b. *Physical*. "My ability in sports."
- c. *Self-confidence*. "I am confident that I can do the job I set out to do."
- d. *Miscellaneous*. "Because he is famous."

A chi-square test was applied to the totals obtained for the three major categories of group behavior, character traits, and achievement for the four groups: male urban, male rural, female urban, and female rural (shown in Table 1). When personality assets are considered by urban and rural males, they perceive themselves differently ($\chi^2 = 35.21$, $df = 2$, $p < .01$). They

TABLE 1
FREQUENCY AND PER CENT BREAKDOWN OF PERSONALITY ASSETS BETWEEN URBAN
AND RURAL HIGH-SCHOOL SENIORS*

Personal asset	Self							
	Male (N = 243)				Female (N = 240)			
	Urban f	%	Rural f	%	Urban f	%	Rural f	%
I. Group behavior								
a. Getting along with others	24	18	7	6	31	24	18	16
b. Adjust to the environment	5	4	2	2	5	4	5	4
c. Concern for people	5	3	2	2	4	3	5	4
d. Friendship	9	7	—	—	11	9	—	—
Total	43	32	11	10	51	40	28	24
II. Character traits								
a. Striving	12	9	4	4	11	9	—	—
b. Character strength	24	18	9	8	17	13	11	10
Total	36	27	13	12	28	22	11	10
III. Achievement								
a. Intellectual	13	10	11	10	13	10	—	—
b. Physical	30	23	75	68	26	20	73	66
c. Self-confidence	11	8	—	—	10	8	—	—
d. Miscellaneous	—	—	—	—	—	—	—	—
Total	54	41	86	78	49	38	73	66
Grand total	133	100	110	100	128	100	112	100

* Boys (self), $\chi^2 = 35.21$, $p < .01$, with 2 *df*.
Girls (self), $\chi^2 = 17.84$, $p < .01$, with 2 *df*.

TABLE 1 (continued)
FREQUENCY AND PER CENT BREAKDOWN OF PERSONALITY ASSETS BETWEEN URBAN
AND RURAL HIGH-SCHOOL SENIORS*

Personal asset	Male (N = 222)				Female (N = 219)			
	Urban		Rural		Urban		Rural	
	f	%	f	%	f	%	f	%
I. Group behavior								
a. Getting along with others	22	19	10	9	26	24	16	15
b. Adjust to the environment	9	8	—	—	—	—	—	—
c. Concern for people	38	33	28	24	6	6	—	—
d. Friendship	2	2	—	—	36	34	25	24
Total	71	62	38	33	68	64	41	39
II. Character traits								
a. Striving	6	5	5	4	7	7	10	10
b. Character strength	20	17	33	29	13	12	19	18
Total	26	22	38	33	20	19	29	28
III. Achievement								
a. Intellectual	7	6	6	5	6	6	—	—
b. Physical	8	7	33	29	11	10	35	33
c. Self-confidence	2	2	—	—	1	1	—	—
d. Miscellaneous	1	1	—	—	—	—	—	—
Total	18	16	39	34	18	17	35	33
Grand total	115	100	115	100	106	100	105	100

* Boys (fathers), $\chi^2 = 21.38$, $p < .01$, with 2 *df*.
Girls (fathers), $\chi^2 = 13.79$, $p < .01$, with 2 *df*.

also perceive their fathers differently ($\chi^2 = 21.38$, $df = 2$, $p < .01$). In the boys' perception of the assets of their mothers, however, no significant difference emerged ($\chi^2 = 4.14$, $df = 2$, $p > .10$).

When the personality assets were judged by urban and rural females, significant differences were found in the perception of themselves ($\chi^2 = 17.84$, $df = 2$, $p < .01$), their fathers ($\chi^2 = 13.79$, $df = 2$, $p < .01$), and their mothers ($\chi^2 = 13.03$, $df = 1$, $p < .01$).

Rural boys, in comparison with urban boys, appeared to attach greater importance to character traits, as assets of their fathers (33 per cent and 22 per cent respectively). Rural boys, furthermore, indicated greater concern with physical attributes of their fathers (29 per cent and seven per cent respectively). Both groups perceived equally the intellectual achievement of their fathers to the same small extent (six per cent).

No significant difference was demonstrated between rural and urban boys in the perception of their mothers' assets. Urban boys did perceive their mothers as more sensitive to group behavior, while rural boys specified their

TABLE 1 (continued)
FREQUENCY AND PER CENT BREAKDOWN OF PERSONALITY ASSETS BETWEEN URBAN
AND RURAL HIGH-SCHOOL SENIORS*

Personal asset	Mother							
	Male (N = 225)				Female (N = 220)			
	Urban f	%	Rural f	%	Urban f	%	Rural f	%
I. Group behavior								
a. Getting along with others	26	22	6	6	29	25	9	9
b. Adjust to the environment	9	8	4	4	5	4	—	—
c. Concern for people	36	31	42	39	54	47	47	45
d. Friendship	1	1	—	—	—	—	—	—
Total	72	62	52	49	88	76	56	54
II. Character traits								
a. Striving	6	5	6	6	1	1	—	—
b. Character strength	20	17	28	26	2	2	17	16
Total	26	22	34	32	3	3	17	16
III. Achievement								
a. Intellectual	5	4	3	3	—	—	—	—
b. Physical	13	11	19	16	23	20	32	30
c. Self-confidence	1	1	—	—	1	1	—	—
d. Miscellaneous	—	—	—	—	—	—	—	—
Total	19	16	22	19	24	21	32	30
Grand total	117	100	108	100	115	100	105	100

* Boys (mothers), $\chi^2 = 4.14$, $p < .20$, with 2 *df*.

Girls (mothers), $\chi^2 = 13.03$, $p < .01$, with 1 *df*.

mothers' maintaining greater character strength. Clearly distinguishing from rural youth, urban boys and girls reported their mothers much more concerned with "getting along with others" (22 per cent and 25 per cent and six per cent and nine per cent respectively) and yet much less directed toward character traits (17 per cent and two per cent and 26 per cent and 16 per cent respectively).

One of the most striking findings contained in Table 1 is that rural girls placed *absolutely no value* on either their own intellectual assets or those of their fathers or their mothers. This finding increases in psychological significance when Table 2 (liabilities) is examined. Rural girls placed no onus on any intellectual inadequacy of their fathers and mothers and only a total value of 14 per cent as a fear of their own inadequacy. When we inspect the table referring to the choices for post high-school direction [see Table 4], surprisingly we note that 15 per cent of rural girls plan on a four-year college program and 25 per cent intend a two-year business school. Since intelligence is received neither favorably or unfavorably, one wonders if the rural girls

plan posthigh-school education for marriage reasons, for elevating social status, or merely for a means of avoiding rural boys (since only two per cent of the rural boys tested noted any plans to attend college).

Table 1 indicates that a much higher degree of identification existed between urban youth and their peer groups than for either of the two groups in the rural population. The pattern is similarly extended to the perception of the assets of both the fathers and mothers. Thus, what appears of greatest importance in the socialization process of the rural group is strongly reflected in the direction of character traits and achievement. And the same percentage favors what this group of rural students "liked best" in their fathers and mothers. These data are overwhelmingly in accord with Reisman's propositions: namely, that rural youth is more likely to be educated in a milieu of the inner-directed culture in which emphasis is placed on duty, old-

TABLE 2
FREQUENCY AND PER CENT BREAKDOWN OF PERSONALITY LIABILITIES OF URBAN
AND RURAL HIGH-SCHOOL SENIORS*

Personal liability	Male (N = 245)				Self			
	Urban		Rural		Urban		Rural	
	f	%	f	%	f	%	f	%
I. Group behavior								
a. Fear of social situations	2	2	—	—	5	4	4	4
b. Inability to make friends	4	3	—	—	2	2	4	4
Total	6	5	0	0	7	6	8	8
II. Character traits								
a. Selfishness	19	15	20	17	17	15	14	14
b. Bad temper	8	6	11	9	14	12	14	14
c. Impatience	2	2	1	1	5	4	—	—
d. Conceit—arrogance	10	8	14	12	9	8	4	4
e. Frivolity	1	1	—	—	2	2	—	—
f. General character weakness	8	6	8	7	—	—	10	10
Total	48	38	54	46	47	41	42	42
III. Personal								
a. Intellect inadequacy	35	28	18	15	13	11	14	14
b. Physical inadequacy	17	14	35	29	27	25	20	18
c. Independence	2	2	1	1	—	—	—	—
d. Lack of confidence	10	8	6	4	6	5	14	14
e. Anxiety	—	—	—	—	1	1	4	4
f. Moodiness	5	3	1	1	1	1	—	—
g. Financial difficulties	2	2	5	4	13	11	—	—
Total	71	57	66	54	61	54	52	50
Grand total	125	100	120	100	115	100	102	100

* Boys (self), $\chi^2 = 0.09$, N.S., 1 df.
Girls (self), $\chi^2 = 0.31$, N.S., 2 df.

fashioned virtues, and parent-centered ideals and propositions (6). Students within this context would internalize the values and attitudes of their parents and perceive success within the boundaries of the posited conformity.

Urban adolescents, however, would be more sensitive to group control. They would more easily change their opinions or those of their parents in order to be found more acceptable by the "significant others" in their peer groups and social cliques.

The perception of success in future goals should reflect what the group considers important rather than a mere identification with parental "dictum." One of the most striking contrasts is the high percentage of perceived value (68 and 66 per cent for rural males and females respectively) placed in physical achievement. Approval for the perception of parental physical assets (29 per cent rural boys, 33 per cent rural girls for fathers and 16 per cent for

TABLE 2 (continued)
FREQUENCY AND PER CENT BREAKDOWN OF PERSONALITY LIABILITIES OF URBAN
AND RURAL HIGH-SCHOOL SENIORS*

Personal liability	Fathers							
	Male (N = 225)				Female (N = 216)			
	Urban f	%	Rural f	%	Urban f	%	Rural f	%
I. Group behavior								
a. Fear of social situations	5	4	2	2	8	7	4	4
b. Inability to make friends	1	1	—	—	—	—	4	4
Total	6	5	2	2	8	7	8	8
II. Character traits								
a. Selfishness	17	15	13	12	25	23	18	17
b. Bad temper	19	17	23	21	17	16	16	15
c. Impatience	1	1	1	1	1	1	—	—
d. Conceit—arrogance	24	21	28	25	36	33	33	31
e. Frivolity	—	—	—	—	1	1	—	—
f. General character weakness	14	12	3	3	2	2	—	—
Total	75	66	68	62	82	76	67	63
III. Personal								
a. Intellect inadequacy	3	3	3	3	11	9	—	—
b. Physical inadequacy	5	4	30	27	7	6	15	14
c. Independence	10	9	5	4	1	1	3	3
d. Lack of confidence	2	2	—	—	1	1	4	4
e. Anxiety	9	6	—	—	—	—	9	8
f. Moodiness	3	3	2	2	—	—	—	—
g. Financial difficulties	2	2	—	—	—	—	—	—
Total	34	29	40	36	20	17	31	29
Grand total	115	100	110	100	110	100	106	100

* Boys (fathers), $\chi^2 = 1.17$, N.S., 1 df.
Girls (fathers), $\chi^2 = 3.83$, $p < .20$, 2 df.

TABLE 2 (continued)
FREQUENCY AND PER CENT BREAKDOWN OF PERSONALITY LIABILITIES OF URBAN
AND RURAL HIGH-SCHOOL SENIORS*

Personal liability	Mothers							
	Male (N = 228)				Female (N = 210)			
	Urban f	%	Rural f	%	Urban f	%	Rural f	%
I. Group behavior								
a. Fear of social situations	12	10	11	10	3	3	14	13
b. Inability to make friends								
Total	<u>12</u>	<u>10</u>	<u>11</u>	<u>10</u>	<u>1</u>	<u>1</u>	<u>14</u>	<u>13</u>
II. Character traits								
a. Selfishness	14	12	24	21	30	29	7	7
b. Bad temper	15	13	10	9	15	14	7	7
c. Impatience	1	1	—	—	1	1	5	5
d. Conceit—arrogance	12	10	8	7	34	33	26	25
e. Frivolity	—	—	3	3	1	1	—	—
f. General character weakness								
Total	<u>12</u>	<u>10</u>	<u>3</u>	<u>3</u>	<u>2</u>	<u>2</u>	<u>5</u>	<u>5</u>
III. Personal								
a. Intellect inadequacy	54	46	48	43	83	80	50	49
b. Physical inadequacy	14	12	9	8	9	8	—	—
c. Independence	12	11	34	29	7	6	27	26
d. Lack of confidence	6	5	—	—	1	1	—	—
e. Anxiety	8	7	—	—	1	1	—	—
f. Moodiness	4	4	—	—	—	—	14	12
g. Financial difficulties	4	4	3	3	—	—	—	—
Total	<u>1</u>	<u>1</u>	<u>8</u>	<u>7</u>	<u>—</u>	<u>—</u>	<u>—</u>	<u>—</u>
Grand total	115	100	113	100	105	100	105	100

* Boys (mothers), $\chi^2 = 0.64$, N.S., 2 df.

Girls (mothers), $\chi^2 = 12.46$, $p < .01$, 1 df.

rural boys and 30 per cent rural girls for mothers) compared with other attributes, the rural sample appeared highly influenced by physical prowess and physical attractiveness to such a degree that the rural females were completely unconcerned with any kind of intellectual assets either for themselves or their mothers.

Of course, what the high influence of physical prowess and attractiveness may actually reflect is the great emphasis given by the adolescent in our culture on physical beauty. Perhaps the present rural sample placed greater emphasis on physical beauty than do urban females because the deal of "the all-American normal" in terms of beauty presented on the screen and TV advertisements is that beauty per se will ensure both personal happiness and success in marriage. Little premium was placed on academic success,

college degrees, or long-range career plans. This has been well documented in studies spanning a generation of adolescent desires and aspirations (3). Although the urban seniors also listed physical attributes as worthwhile indicators of their own assets (23 and 20 per cent for males and females respectively), they did indicate that their own assets were also distributed among factors of intelligence, self-confidence, and peer-group acceptance.

2. Liabilities

The 1,341 personality liabilities listed by the subjects were also classified into three major groups: group behavior, character traits, and personal, with each group further divided into subgroups adapted from Reisman's questionnaire (7) as modified by Gaier and Wambach (4). Group behavior, specifically mentioned as detrimental to success, contained some of the following representative responses:

- a. *Fear of social situations.* "She's afraid to go places where there are crowds of people."
- b. *Inability to make friends.* "Some people are hard for me to be friendly with."

The subcategories of character traits included the following:

- a. *Selfishness.* "I always want my own way."
- b. *Bad Temper.* "He gets awful mad."
- c. *Impatience.* "He gets upset with people."
- d. *Conceit and arrogance.* "He thinks he is always right."
- e. *Frivolity.* "I'm not serious enough; everything is a big joke."
- f. *General Character Weakness.* "He's not much good."

The personal category which provides an index of the inadequacy the individual perceives for himself or his parents in the challenge for a successful life contained the following subdivisions:

- a. *Intellectual inadequacy.* "My school marks are pretty bad."
- b. *Physical inadequacy.* "My hips." "My appearance."
- c. *Independence.* "I don't like to be told what to do."
- d. *Lack of confidence.* "He always thinks he is inferior."
- e. *Anxiety.* "She's always worried about something."
- f. *Moodiness.* "Her mood changes just like that."
- g. *Financial difficulties.* "He can't handle money."

A chi-square test was applied to the number of the dispersion among liabilities [see Table 2]. Only one significant difference in any area emerged: girls in the perception of their mothers significant at the .01 level. The girls' perception of the liabilities of their mothers is heightened by the lack of sig-

nificance in the other two categories. In 1950, Stokes (9) hypothesized that girls in our cultural context have much more behavioral identification with their mothers than boys with their fathers. What might well be reflected in our data on liabilities is an index of behavioral identification rather than emotional identification. Urban females attributed 80 per cent of their mothers' liabilities to character traits; only eight per cent of the liabilities was founded on intellectual inadequacy. Were these urban girls expressing a behavioral identification in society but not an emotional behavior because they did not wish to identify with the "bad character" of their mothers?

It is especially noteworthy that neither urban nor rural students perceived lack of specific abilities and fears of group behavior as meaningful liabilities within their social cultures, whereas group behavior characteristics were significant assets in "belonging" to their respective adolescent societies. Among the personality liabilities, "selfishness" and "conceit-arrogance" were of primary concern for the students in both schools but to a marked degree for urban girls.

Urban girls were consistently adamant about these two liabilities in their mothers (29 per cent and 33 per cent of their listings for "selfishness" and "conceit-arrogance" respectively).

Rural boys were not as concerned about their own physical limitations and liabilities as were urban youth; rather, they were considerably more sensitive to the physical "defects" of both parents than were urban boys and girls. The highly prestigious value of physical assets, which rural youth demand for recognition in adolescent society, diminished within the scale of comparable liabilities. Even though the boys and girls of urban culture appeared more concerned than rural youth about their *own* lack of achievement, nevertheless, rural youth perceived a deficiency of personal achievement among their parents as more unfavorable than their urban counterparts.

3. *Postgraduate Plans*

As part of the questionnaire, each subject was asked to list the occupation of each of his parents, as well as his own future occupational and educational plans. These were then classified using an adaptation of the system reported by Rogoff (8) in which she developed a measurement of occupational ability within each stratum of society. The per cent breakdown of the parents' occupation is shown in Table 3; Table 4 lists the expected vocational aspirations of the present subjects.

When the plans of high-school seniors are compared with the occupations of their parents according to urban and rural orientation, expected significant

TABLE 3
PER CENT BREAKDOWN OF PARENT'S CURRENT OCCUPATIONS (TOTAL $N = 170$)

	Urban (N = 86)						Rural (N = 84)					
	Boys (N = 53)			Girls (N = 33)			Boys (N = 44)			Girls (N = 40)		
Parent's occupation	Fathers	Mothers		Fathers	Mothers		Fathers	Mothers		Fathers	Mothers	
1. Professional and semiprofessional	11	8		18	6		2	—		10	5	
2. Proprietors, managers, and officials	13	2		39	—		7	—		20	—	
3. Clerical and sales	8	24		7	26		2	27		5	19	
4. Skilled	11	—		6	—		9	—		5	—	
5. Semiskilled	34	2		18	6		22	9		3	—	
6. Protective service	2	—		3	—		—	—		—	5	
7. Other service, personal, unskilled farming, etc.	3	2		—	3		60	9		42	16	
8. Retired, deceased, not working	18	62		9	59		18	55		15	55	
	100	100		100	100		100	100		100	100	

TABLE 4
PER CENT BREAKDOWN OF POSTHIGH SCHOOL AND OCCUPATION CHOICES

Choice	Urban		Rural	
	Boys (N = 53)	Girls (N = 30)	Boys (N = 44)	Girls (N = 40)
Four-year college	54	43	2	15
Technical or business school	23	20	18	25
Specific skill, labor, trade	—	9	10	10
Chemical, general labor, sales, army, farming	19	22	50	15
Don't know. Anything reasonable.	4	6	20	35
Total	100	100	100	100

differences in the two cultures emerge. Urban adolescents are considerably more concerned about formal schooling upon completion of high school than rural seniors. Whereas 54 per cent of the urban males and 43 per cent of urban females planned to seek admission to a four-year college, only two per cent of the males and 15 per cent of the females in the rural schools were considering college. Technical or business school attendance was planned by 23 per cent of the males in the city school and 18 per cent of the males in the rural community. A greater percentage of rural girls desired business-school courses than urban girls. Half of the rural sample of males would seek farming, the military service, personal jobs, or unskilled employment; but only 19 per cent of the urban boys stated these as work choices. The plans of the urban youth were more specific. Only four per cent of the urban males and six per cent of the urban girls were undecided, while 20 per cent of the rural boys had not made up their minds and 35 per cent of the girls would do "something that is reasonable."

From Table 3, we can observe that the occupation of the fathers of the rural group was heavily directed to farming and unskilled labor (60 per cent and 52 per cent respectively). Two per cent of the city labor group had fathers who were unskilled laborers. As expected, urban fathers were more heavily engaged in professional and semiprofessional, managerial, and sales positions.

E. CONCLUSION

In line with Reisman's assessment of the orientation of subjects according to an inner-directed and other-directed personality hypothesis, the present rural subjects appear to follow more frequently the old-fashioned virtues, the principles, and philosophy of their parents. By contrast, urban youth apparently were more directed by group trends. This is not to suggest that these rural youth were inner-directed isolates. While the individual may appear to be

influenced by varying aspects of tradition, inner-directed, and other-directed culture, Reisman's theory holds that one dominant socializing agent pervades a given personality. Thus, while there is much overlapping among the categories, each individual lives in tune with the operative function of his culture.

In the present study, the rural sample has emerged as more inner directed, with the subsequent choosing occupations and careers in accord with the occupations and careers of their parents. In a sense, they have rejected an interest in continuing formal education; in the same manner, their fathers and mothers have not fulfilled themselves in this way. Urban youth, on the contrary, perceived themselves as sensitive to group thinking and behavior. Since they are living in an environment in which their peers deem continuing education as an important means of mobility, these youth are more concerned with college education. And it does seem apparent that both peer groups and parents are influencing urban high-school seniors to advance up the educational ladder as a basis for growth toward personal self-actualization as well as mobility.

REFERENCES

1. COLEMAN, J. *The Adolescent Society*. Glencoe, Ill.: Free Press, 1961.
2. ESKOLA, A. Level of aspiration in career selection. *Acta Sociologica*, 1962, 5, 238-248.
3. GAIER, E. L., & LITTUNEN, Y. Modes of conformity in two subcultures: A Finnish American comparison. *Acta Sociologica*, 1961, 5, 65-75.
4. GAIER, E. L., & WAMBACH, H. S. Self-evaluation of personality assets of southern White and Negro students. *J. Soc. Psychol.*, 1960, 51, 135-143.
5. LITTUNEN, Y., & GAIER, E. L. Occupational values and modes of conformity: Turku, Finland. *J. Soc. Psychol.*, 1960, 51, 123-133.
6. REISMAN, D. *The Lonely Crowd*. New Haven, Conn.: Yale Univ. Press, 1950.
7. ———. *Faces in the Crowd*. New Haven, Conn.: Yale Univ. Press, 1952.
8. ROGOFF, N. *Occupational Mobility*. Glencoe, Ill.: Free Press, 1953.
9. STOKES, S. An inquiry into the concept of identification. *J. Gen. Psychol.*, 1950, 76, 163-189.

State University of New York at Buffalo
Buffalo 14, New York

BOOKS RECENTLY RECEIVED

- AICHHORN, A. *Delinquency and Child Guidance*. New York: International Univ. Press, 1964. Pp. 244.
- ALLPORT, G. W., *Ed.* *Letters from Jenny*. New York: Harcourt, Brace & World, 1965. Pp. 223.
- AMOS, JIMMY R., BROWN, F. L., & MINK, O. G. *Statistical Concepts: A Basic Program*. New York: Harper & Row, 1965. Pp. 125.
- BECK, S. J. *Psychological Processes in the Schizophrenic Adaptation*. New York: Grune & Stratton, 1965. Pp. 421.
- BERNDT, R. M., & BERNDT, C. H. *The World of the First Australians*. Chicago: Univ. Chicago Press, 1964. Pp. 509.
- BEST, H. *Public Provision for the Mentally Retarded in the United States*. New York: Thomas Y. Crowell, 1965. Pp. 455.
- CANTOR, A. J. *Unitrol: The Healing Magic of the Mind*. West Nyack, N. Y.: Parker, 1965. Pp. 192.
- CLARK, K. B. *Dark Ghetto*. New York: Harper & Row, 1965. Pp. 251.
- DENNIS, W. *The Hopi Child*. New York: John Wiley & Sons, 1965. Pp. 200.
- DUBOIS, P. H. *An Introduction to Psychological Statistics*. New York: Harper & Row, 1965. Pp. 530.
- DUNNETTE, M. D., & KIRCHNER, W. K. *Psychology Applied to Industry*. New York: Appleton-Century-Crofts, 1965. Pp. 262.
- EBIN, D., *Ed.* *The Drug Experience*. New York: Grove Press, 1965. Pp. 385.
- GROSSMAN, H. *Teaching the Emotionally Disturbed*. New York: Holt, Rinehart & Winston, 1965. Pp. 184.
- HAMACHEK, D. E. *The Self in Growth, Teaching, and Learning*. Englewood Cliffs, N. J.: Prentice-Hall, 1965. Pp. 576.
- HARTLEY, R. E., & HARTLEY, E. L. *Readings in Psychology* (3rd ed.) New York: Thomas Y. Crowell, 1965. Pp. 612.
- ISAACSON, R. L., HUTT, M. L., & BLUM, M. L. *New York: Harper & Row, 1965. Pp. 374.*
- KAPLAN, L. *Foundations of Human Behavior*. New York: Harper & Row, 1965. Pp. 368.
- KESSEN, W. *The Child*. New York: John Wiley, 1965. Pp. 301.
- KLINEBERG, O., & CHRISTIE, R., *Eds.* *Perspectives in Social Psychology*. New York: Holt, Rinehart & Winston, 1965. Pp. 214.
- KRASNER, L., & ULLMAN, L. P. *Research in Behavior Modification*. New York: Holt, Rinehart & Winston, 1965. Pp. 403.
- LOREE, M. R. *Psychology of Education*. New York: Ronald Press, 1965. Pp. 620.
- LYONS, J. *A Primer of Experimental Psychology*. New York: Harper & Row, 1965. Pp. 322.
- MEYER, H. J., BORGATTA, E. F., & JONES, W. C. *Girls at Vocational High*. New York: Russell Sage, 1965. Pp. 225.
- MONEY, J., *Ed.* *Sex Research*. New York: Holt, Rinehart & Winston, 1965. Pp. 260.
- MUELLER, C. G. *Sensory Psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1965. Pp. 120.
- MUSGROVE, F. *Youth and the Social Order*. Bloomington: Indiana Univ. Press, 1965. Pp. 168.
- NG, L. K. Y., & MUDD, S., *Eds.* *The Population Crisis*. Bloomington: Indiana Univ. Press, 1965. Pp. 364.
- OSGOOD, C. E., & SEBEOK, T. A., *Eds.* *Psycholinguistics: A Survey of Theory and Research Problems*. Bloomington: Indiana Univ. Press, 1965. Pp. 307.

- PETRIE, S. How to Reduce and Control your Weight through Self-Hypnotism. Englewood Cliffs, N. J.: Prentice-Hall, 1965. Pp. 223.
- PINCKNEY, E. R., & PINCKNEY, C. The Fallacy of Freud and Psychoanalysis. Englewood Cliffs, N. J.: Prentice-Hall, 1965. Pp. 186.
- PROSHANSKY, H., & SEIDENBERG, B., *Eds.* Basic Studies in Social Psychology. New York: Holt, Rinehart & Winston, 1965. Pp. 742.
- RICE, A. K. Learning for Leadership. New York: Humanities Press, 1965. Pp. 200.
- ROACH, M. E., & EICHER, J. B. Dress, Adornment and the Social Order. New York: John Wiley, 1965. Pp. 429.
- ROGLER, L. H., & HOLLINGSHEAD, A. B. Trapped Families and Schizophrenia. New York: John Wiley, 1965. Pp. 436.
- SAMPSON, R. C. Managing the Managers. New York: McGraw-Hill, 1965. Pp. 272.
- SCHEIN, E. H., & BENNIS, W. G. Personal and Organizational Change through Group Methods: The Laboratory Approach. New York: John Wiley, 1965. Pp. 376.
- SCHILDER, P. Medical Psychology. New York: John Wiley, 1965. Pp. 428.
- SEARLES, H. F. Collected Papers on Schizophrenia and Related Subjects. New York: International Univ. Press, 1965. Pp. 797.
- STEINER, G. A., *Ed.* The Creative Organization. Chicago: Univ. Chicago Press, 1965. Pp. 267.
- STEINER, I. D., & FISHBEIN, M., *Eds.* Current Studies in Social Psychology. New York: Holt, Rinehart & Winston, 1965. Pp. 532.
- STOTLAND, E., & KOBLER, A. L. Life and Death of a Mental Hospital. Seattle: Univ. Washington Press, 1965. Pp. 257.
- TIFFIN, J., & MCCORMICK, E. J. Industrial Psychology (5th ed.) Englewood Cliffs, N. J.: Prentice-Hall, 1965. Pp. 682.
- TOMKINS, S. S., & IZARD, C. E., *Eds.* Affect, Cognition and Personality. New York: Springer, 1965. Pp. 464.
- TRIST, E. L., HIGGIN, G. W., MURRAY, H., & POLLOCK, A. B. Organizational Choice. New York: Humanities Press, 1965. Pp. 332.
- ULLMAN, L. P., & KRASNER, L., *Eds.* Case Studies in Behavior Modification. New York: Holt, Rinehart & Winston, 1965. Pp. 396.
- WATSON, R. I. Psychology of the Child (2nd ed.) New York: John Wiley, 1965. Pp. 635.
- YOUNG, F. W. Initiation Ceremonies: A Cross-Cultural Study of Status Dramatization. New York: Bobbs-Merrill, 1965. Pp. 199.
- YOUNG, R. K., & VELDMAN, D. J. Introductory Statistics for the Behavioral Sciences. New York: Holt, Rinehart & Winston, 1965. Pp. 435.
- ZUBIN, J., ERON, L. D., & SCHUMER, F. An Experimental Approach to Projective Techniques. New York: John Wiley, 1965. Pp. 645.